



Séquence d'images et techniques de codage MICDA intra/interimage. Rapport de synthèse

Claude Labit

► To cite this version:

Claude Labit. Séquence d'images et techniques de codage MICDA intra/interimage. Rapport de synthèse. [Rapport de recherche] RR-0630, INRIA. 1987. inria-00075923

HAL Id: inria-00075923

<https://inria.hal.science/inria-00075923>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNITÉ DE RECHERCHE
INRIA-RENNES

Rapports de Recherche

N° 630

**SÉQUENCE D'IMAGES
ET
TECHNIQUES DE CODAGE
MICDA INTRA/INTERIMAGE
RAPPORT DE SYNTHÈSE**

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél: (1) 39 63 55 11

Claude LABIT

Février 1987

Campus Universitaire de Beaulieu
35042 - RENNES CÉDEX
FRANCE
Téléphone: 99 36 20 00
Télex: UNIRISA 950 473 F
Télécopie: 99 38 38 32

Publication Interne n°339

Janvier 1987 - 140 pages

SEQUENCE D'IMAGES ET TECHNIQUES DE CODAGE MICDA

INTRA/INTERIMAGE:

RAPPORT DE SYNTHESE

Claude LABIT

Résumé :

Ce rapport présente l'approche suivie pour l'ensemble de nos recherches et expérimentations réalisées dans le domaine du codage de séquences d'images de télévision. Les techniques employées concernent les schémas de codage adaptatifs tant dans le domaine spatial (intraimage) que dans le domaine spatiotemporel (interimage) avec estimation de mouvement. Ce rapport donne les principaux résultats fournis par de tels algorithmes concernant la réduction de débit de transmission obtenu et la robustesse de ceux-ci en présence d'erreurs de transmission sur un canal. En conclusion, l'approche MICDA est comparée à d'autres perspectives d'études qui sont en cours actuellement.

IMAGE SEQUENCES AND ADPCM INTRA-INTERIMAGE : CODING TECHNIQUES FINAL REPORT

Abstract :

We present a review of our simulations and experimental results that we did, during all the four last years, about television sequence image coding. The followed approach uses adaptive coding DPCM schemes applied simultaneous to spatial domain and to spatiotemporal domain with motion estimation algorithms. This paper gives the main results obtained by these techniques and evaluates bit rate reduction and error traps mission robustness of the different coding schemes. In conclusion, ADPCM techniques are compared to other methods that we are currently testing.



**ETUDE DE SCHEMAS DE CODEURS MICDA INTRA/INTERIMAGES
POUR LES SEQUENCES D'IMAGES DE TELEVISION**

CHAPITRE I : INTRODUCTION	3
I.1 - Cadre des études	3
I.2 - Collaboration de recherche CCETT/INRIA	4
I.3 - Enchaînement des études algorithmiques	5
I.4 - Nature des travaux	7
I.5 - Présentation des chapitres suivants	13
 CHAPITRE II : RAPPEL HISTORIQUE ET RESUME DES DIFFERENTES PHASES D'ETUDES	 15
II.0 - Le pourquoi du A(DPCM)	15
II.1 - Historique n°1 : Simulation d'un codeur adaptatif intraimage	17
II.2 - Historique n°2 : Réalisation hardware d'un codeur adaptatif intraimage	20
II.3 - Historique n°3 : Etude mouvement	24
II.4 - Historique n°4 : Schémas de codage multimodes	29
II.5 - Conclusion partielle	35
 CHAPITRE III : EXPERIMENTATIONS COMPLEMENTAIRES	 37
III.1 - Simulations sur séquences "cost"	37
III.2 - Robustesse d'algorithmes adaptatifs multimodes intra/interimages aux erreurs de transmission	58
 CONCLUSIONS ET PERSPECTIVES	 71
 REFERENCES BIBLIOGRAPHIQUES	 75
 ANNEXES	
Annexe I	79
Annexe II	81
Annexe III	91
Annexe IV	101
Annexe V	105
Annexe VI	109
Annexe VII	111
Annexe VIII	121
Annexe IX	125

Préambule

Ce document a pour objectif de réunir et de synthétiser au sein d'une même chronologie les différentes simulations (et résultats algorithmiques associés) effectuées pendant la période 1984-1986. Ces résultats ont été exposés et publiés de ci de là mais nécessitent toutefois une présentation ordonnée et argumentée. Tel est le sujet de ce rapport.

Il nous est apparu nécessaire d'y inclure d'importants rappels historiques sur l'ensemble des recherches que nous avons menées tant dans le domaine intraimage que interimage.

Nous tenons à mentionner le fait que ces études ont été menées dans le cadre d'une étroite collaboration entre l'IRISA et le CCETT (Département RTI : études codage) et ont été réalisées tant dans le cadre de convention d'études (ce rapport final de convention en est l'illustration) que d'actions de recherche "bénévoles" et spontanées, instructives pour tous. Nous tenons à ce titre à remercier l'ensemble des acteurs ayant contribué à ces travaux.

CHAPITRE I

INTRODUCTION

I.1 - CADRE DES ETUDES

Le cadre général de ces études est la conception de schémas de codage (-décodage) "hautement adaptatifs" pour transmettre des séquences d'images de Télévision avec une réduction maximale du débit de transmission tout en conservant une haute qualité dans les images transmises (qualité "broadcast").

Le cadre algorithmique retenu était le codage MICDA (codage différentiel adaptatif) utilisant les **corrélations spatiotemporelles** résidentes au sein d'une séquence d'images pour extraire l'innovation résiduelle d'une image par prédiction et différenciation entre l'image originale et une prédiction adaptative spatiotemporelle de celle-ci. L'optimisation de techniques visant à réduire le débit de transmission en sortie du codeur était donc à rechercher selon deux directions:

- *recherche de schémas de prédiction spatiotemporels adaptatifs "optimaux",
- *recherche de schémas de "compactage" de l'information à transmettre.

Nous avons ainsi intitulé ces études : conception de schémas de codage A-DPCM hybride intra/interimage avec compensation de mouvement; en effet elles ont réuni:

- le savoir faire acquis dans le domaine du codage ADPCM intraimage (études antérieures à cette convention)
- le savoir faire acquis dans le domaine de l'estimation de mouvement et son application au codage interimage avec compensation de mouvement

- la définition de schémas de prédiction mixte intégrant à la fois les schémas de prédiction intraimage et interimage. C'est le sens que nous donnons au terme "hybride".

- la définition de schémas de codage complets (avec évaluation du débit de transmission associée et de la robustesse algorithmique aux erreurs de transmission).

I.2 - COLLABORATION DE RECHERCHE CCETT/INRIA Rennes

Ce document constitue le rapport final concluant la convention de recherche n°82 1B 146 00-790-92-45 entre le CCETT et l'INRIA (Centre de Rennes). L'annexe technique précise les contraintes qui avaient été posées en préalables à l'étude:

***Caractéristiques sur les séquences d'images-source à coder**

- haute qualité d'image
- type de scène quelconque
- type de mouvement présent dans la séquence quelconque.

***Contraintes sur l'algorithmique à développer**

- respect de la qualité des images en reconstruction
- robustesse vis à vis des erreurs de transmission
- pas d'excès de complexité algorithmique dans un souci de réalisme "hardware".

Le but de ce rapport est également de regarder dans quelles mesures ces contraintes ont été traitées et/ou respectées. La démarche générale pour mener cet étude a toutefois été de réaliser une étude de recherche "amont" visant à extraire de manière maximale l'ensemble des corrélations spatiotemporelles résidentes au sein de telles séquences d'images.

I.3 - ENCHAINEMENT DES ETUDES ALGORITHMIQUES

La figure 1 jointe donne l'enchaînement temporel des différentes phases d'étude.

*partant d'une base de données d'images fixes (images fixes standards: "bateau", "batiment"...), l'introduction d'adaptativité a tout d'abord porté sur le choix de prédicteurs spatiaux adaptés aux configurations locales du signal d'image (fixe) dans le cadre de schéma de codage "intraimage".

*A partir de séquences "simples" d'images visiophoniques (Judith) et de scènes en mouvement de maquettes de jouets ("chateau"), de nombreuses études comparatives ont été faites sur différents schémas d'estimation de mouvement en vue d'un système de codage adaptatif avec compensation de mouvement.

*Enfin, à partir des séquences standard (normalisées au niveau cost 211 bis européen "voiture", "girl" et "musi") un schéma de codage hybride intra/interimage avec compensation de mouvement a été simulé et testé ainsi que sa robustesse vis à vis des erreurs de transmission. La séquence "voiture", pour ce faire, a été plus particulièrement utilisée pour ses caractéristiques critiques:

- quant au contenu de scène (nombreux objets différents et critiques tels que la grille,...)
- quant à son important taux de mouvement (objets mobiles et caméra en mouvement)
- quant à son rapport signal/bruit défavorable.

A l'issue de chacune de ces phases d'étude, des conclusions partielles ont été énoncées. Nous en rappelons brièvement quelques unes d'entre elles dans le chapitre suivant. Les références correspondantes existent pour une lecture plus détaillée.

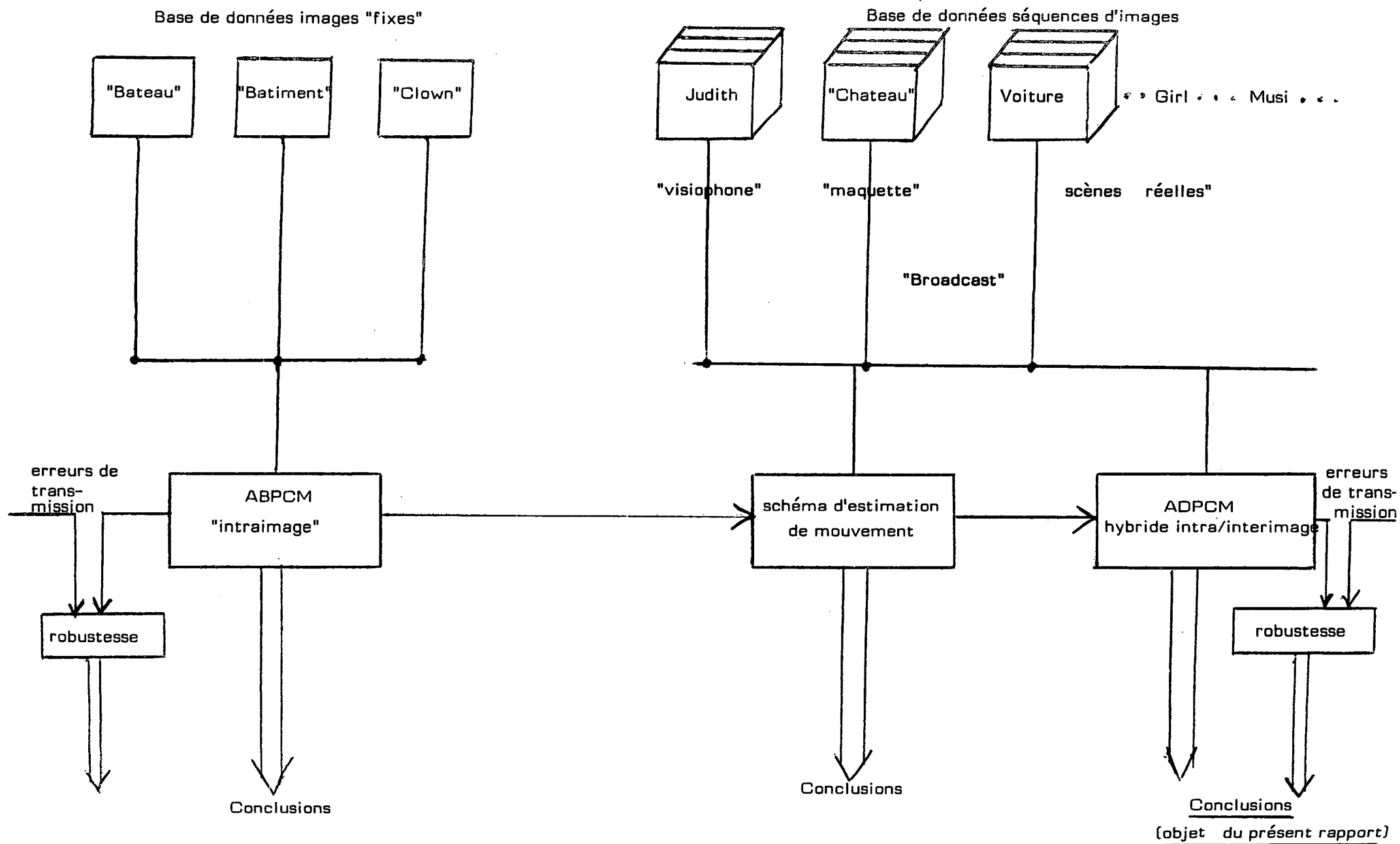


Figure 1 : Enchaînement des études

I.4 - NATURE DES TRAVAUX

I.4.1 - Approche locale - Approche globale ?

On se reportera utilement aux articles généraux [12] et aux publications récentes pour situer le contexte de ces études concernant les schémas de codage prédictifs adaptatifs. Rappelons brièvement que l'existence au sein d'une séquence d'images de fortes corrélations spatiales, temporelles et spatiotemporelles conduit naturellement à la conception de schéma de codage utilisant ces différents domaines pour bâtir des prédicteurs adaptés (au sens théorie de l'information et perception visuelle des dégradations). L'adaptatif peut être introduit à différents niveaux hiérarchiques dans l'image, par exemple:

- algorithmes à stratégies d'estimation et de prédiction par blocs
- algorithmes à stratégies fortement locales.

Nous avons opté dans le cadre de ce travail pour l'approche fortement locale visant à définir **en chaque pixel** une prédiction "optimale" (choisie parmi une batterie de prédicteurs potentiels prédéfinis) et donc une erreur de prédiction en chaque point d'image la plus faible ("ou moins visible") possible. Cette approche fortement locale a toutefois été compensée par l'étude de schémas de segmentation de l'image et de compactage de l'information à transmettre la moins coûteuse en terme de débit de transmission.

Si $I(x,t)$ désigne la séquence d'images étudiée (seule l'information de luminance est abordée ici) et $C(x,t)$ un vecteur caractérisant le pixel (x,t) et contenant des informations qualitatives (point "contour", point homogène, point texture, point fixe, point en mouvement...) et/ou quantitatives (contour d'orientation $C(x,t) = \theta$, point

en mouvement $C(x,t) = \hat{d}(x,t)$ vecteur mouvement estimé en ce point...) alors il convient de trouver la loi de prédiction locale f définie par:

$$P(x,t) = I(x,t) = f(C(x,t), V_I^{\mathcal{V}}(x,t))$$

avec

$\hat{I}(x,t)$: luminance prédite au point (x,t)

$C(x,t)$: caractéristique spatiotemporelle du point (x,t) ,

$V_I^{\mathcal{V}}$: ensemble des luminances reconstruites dans un voisinage spatiotemporel causal du point (x,t) .

L'erreur de prédiction quantifiée (quantificateur Q) est, avec ces notations, définie par :

$$\begin{aligned} \tilde{e}(x,t) &= Q(e(x,t)) = Q(I(x,t) - \hat{I}(x,t)) \\ &= Q(I(x,t) - f(C(x,t), V_I^{\mathcal{V}}(x,t))) \end{aligned}$$

Les luminances reconstruites sont alors définies par :

$$\tilde{I}(x,t) = \tilde{e}(x,t) + f(C(x,t), V_I^{\mathcal{V}}(x,t))$$

Nous donnerons ci-dessous quelques exemples de fonctions locales f expérimentées. Le domaine de définition de cette fonction f peut être restreinte au cas intraimage ou au cas interimage. Les études décrites dans le paragraphe II.5 et le chapitre III porteront sur un domaine de définition 3-D spatiotemporel (x,t) .

I.4.2 - Caractérisation locale ou globale?

En reprenant le formalisme utilisé précédemment, deux approches algorithmiques peuvent être choisies quant à la nature de la fonction $C(x,t)$ de caractérisation.

* $C(x,t)$ peut fournir les paramètres caractéristiques d'un ensemble de pixels c'est à dire être constante par morceaux sur son domaine de définition.

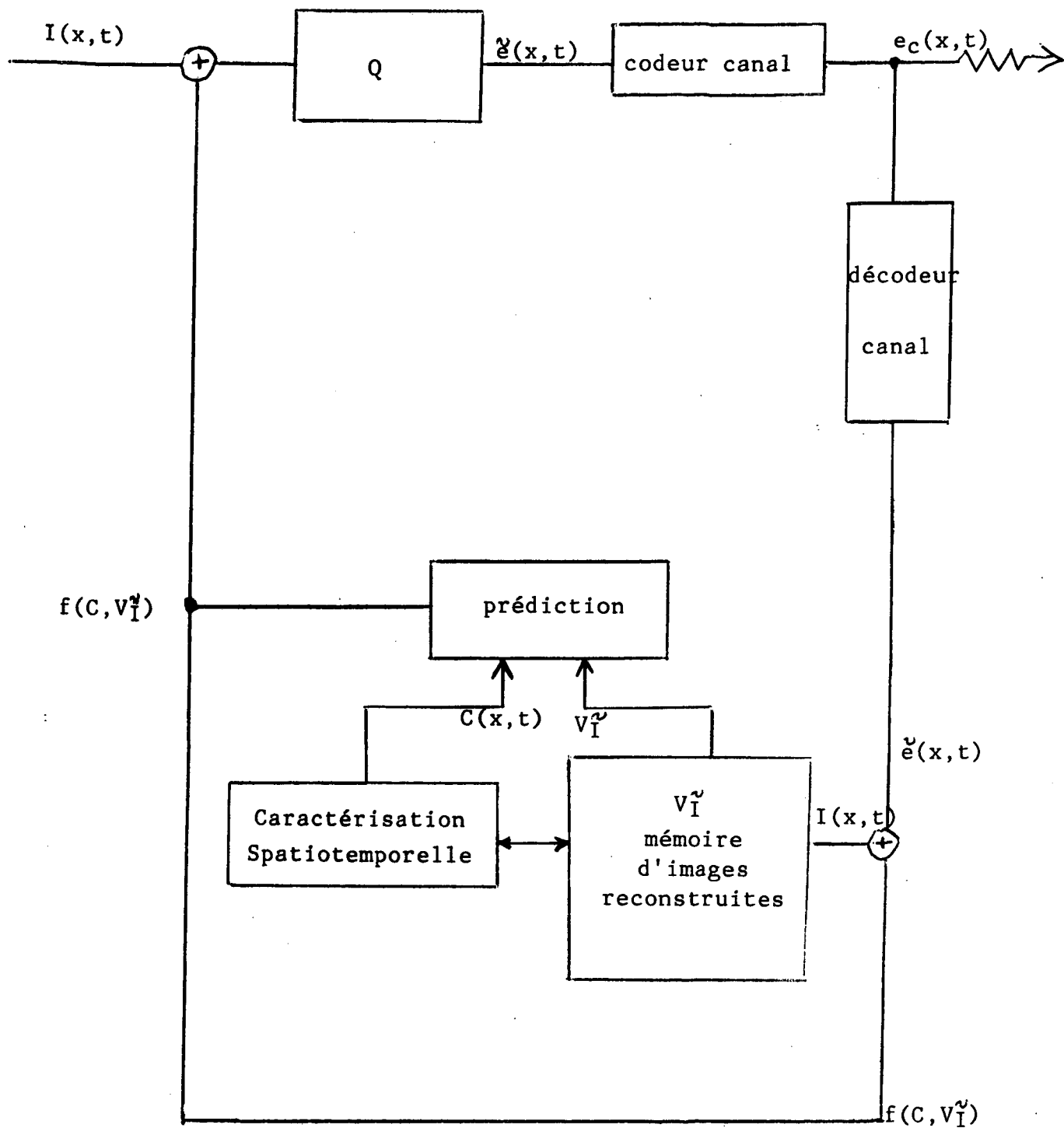


Figure 2
Schéma d'un codeur MICPA (A-DPCM
à caractérisation locale prédicte

* $C(x,t)$ peut varier en chaque point et donc représenter la configuration locale d'un pixel.

Afin de poursuivre notre investigation dans le cadre de la recherche d'algorithme de codage "hautement" adaptatif, nous avons choisi la seconde approche. Nous indiquons dès maintenant cependant que les caractérisations $C(x,t)$ et $C(x',t)$ ne seront pas choisies indépendamment les unes des autres sur des points x et x' voisins afin de compacter au mieux l'information à transmettre.

Si nous prenons l'exemple de la compensation de mouvement pure c'est à dire, où

$$C(x,t) = \hat{d}(x,t) \text{ estimée du vecteur de déplacement au point } (x,t)$$

alors nous retrouvons les deux approches classiques en matière d'algorithmes d'estimation de mouvement:

caractérisation globale de $C(x,t)$	→	estimateur de mouvement par bloc : 13 , 14
--	---	--

caractérisation locale de $C(x,t)$	→	estimateur de mouvement chaque pixel : algo pel-récuratif 5 , 7
---------------------------------------	---	--

I.4.3 - Prédiction ou transmission ?

Le vecteur de caractérisation $C(x,t)$ peut être de deux natures:

*une estimation instantanée au point (x,t) de la configuration locale qui est calculée au codeur et doit être transmise au décodeur ce qui constitue bien souvent une partie non-négligeable du débit à transmettre.

*une prédiction basée sur des informations antérieurement reconstruites (contrainte de causalité spatiotemporelle). Ainsi dans le cadre de la composition de mouvement,

$$C(x,t) = \hat{a}(x,t) = \hat{a}(x-1,t)$$

(x-1) pouvant désigner un des points du voisinage

$$V_I(x,t) \quad \text{Image } (t)$$

soit par exemple le point précédent (x_1-1, y_1)
ou le point de la ligne précédente (x_1, y_1-1)
pour le point $x = (x_1, y_1)$.

Nous avons opté pour ce second cas dans lequel aucune information de caractérisation définissant $C(x,t)$ n'est à transmettre sur le canal. Le codeur et le décodeur réaliseront symétriquement la même phase d'estimation de ces caractéristiques spatiotemporelles locales. Ce choix a été effectué dans l'idée d'opérer une **réduction maximale** du débit de transmission sans envoi "parasite" d'informations redondantes car réestimables au décodeur. Ce choix cependant a des conséquences au niveau de la robustesse du schéma de codage-décodage vis à vis d'erreurs de transmission. Une discussion concernant cette question est faite au chapitre III de ce présent rapport. A contrario, l'autre optique peut être prise et donner des résultats similaires au vue du seul critère de réduction de débit [14].

I.4.4 - Débit fixe ou débit variable ?

L'une des techniques historiquement la plus répandue dans le domaine du codage temporel consiste en le rafraichissement conditionnel ("conditional replesnihment") d'une séquence d'images.

La transmission d'inforamtion est conditionnée à la présence d'innovation temporelle; si celle-ci n'est pas suffisante aucune information n'est transmise sur le canal; le débit résultant est ainsi variable. La même philosophie peut être étendue au cas de codeur où l'envoi d'informations

n'est pas systématique. Dans le même souci (de réduction maximale du débit) que celui mentionné au paragraphe précédent, nous avons opté pour l'approche débit-variable; ainsi l'erreur de prédiction notée $e(x,t)$ au paragraphe I.4.1 n'est pas envoyée systématiquement. Ce choix a des conséquences sur:

- la nature des informations transmises; des informations de synchronisation (adresses de plages) sont à transmettre dans le cas de codage conditionnel.
- la robustesse du schéma de codage générant un train binaire composé d'informations différentes plus ou moins sensibles vis à vis d'apparition d'erreurs de transmission.
- la nécessité de réguler le débit variable en lui apposant la contrainte du débit-fixe du canal. Cette phase d'étude n'a pas été abordée dans le cadre des expérimentations réalisées pour cette convention d'études.

I.4.5 - Résumé et remarques :

La figure 2 résume les choix algorithmiques retenus:

- stratégie de prédiction locale
- schéma de codage à caractéristique locale spatiotemporelle prédictive
- test sur l'erreur de prédiction et transmission non systématique.

Nous résumons ces différents choix, par la notion de schémas de codage "**hautement adaptatifs**" et travaillant dans le domaine hybride **intra/interimage**; cette étude est à ce titre une recherche "amont" qui visait à extraire de manière maximaliste l'ensemble des corrélations spatiotemporelles présentes au sein d'une séquence d'images de type Télévision professionnelle. Le critère pour l'élaboration de ces choix ne fut pas en premier lieu la réduction de débit résultante.

A contrario, en tant qu'évaluations a posteriori, nous avons mesuré le débit de transmission que de tels schémas de codage fournissaient et testé leurs robustesses vis à vis de l'introduction d'erreurs-canal de transmission. Les résultats de ces évaluations sont donnés dans les annexes jointes à ce rapport. Nous les avons brièvement résumés dans le chapitre III.

Une discussion critique et comparative par rapport à d'autres approches prises (cf. bibliographie [10], [14] s'impose suite à cette étude.

I.5 - PRESENTATION DES CHAPITRES SUIVANTS

Le chapitre II sera consacré aux rappels méthodologiques et au résumé des différentes phases d'étude réalisées tant dans le domaine du codage MICDA intraimage que tant dans l'exploration des diverses techniques d'estimation de mouvement. Les principaux résultats et conclusions de ces phases d'études seront données brièvement. Ces rappels permettront de décrire avec précision le système de codage multimode intra/interimage qui a été expérimenté. De nombreuses variations de cet algorithme ont été testé et simulé comparativement ([10]) sur plusieurs types de données de séquences d'images. Seule la version définitive sera rappelée et détaillée ici dans un souci de synthèse.

Le chapitre III fournira les principaux enseignements tirés suite:

- à l'évaluation sur séquences COST 211 bis (voiture, girl, musi) de cet algorithme de codage,
- à l'étude de la robustesse de celui-ci vis à vis des erreurs de transmission.

Les résultats associés à ces expérimentations ([7] à [10] et Annexes) ont d'ores et déjà été exposés et/ou publiés mais dans des contextes dispersés et peu homogènes. Le chapitre III vise ainsi à donner un aperçu condensé de l'ensemble de ces études.

Nous souhaitons conclure au chapitre IV par un bilan concernant l'ensemble de ces études, notamment en discutant des choix algorithmiques pris a priori et en comparant nos résultats à d'autres résultats publiés dans la littérature concernant les codeurs A-DPCM avec compensation de mouvement. Ces conclusions seront à reliées également aux perspectives d'études que nous avons d'ores et déjà entreprises dans un tout autre domaine, le codage spatiotemporel par transformées.

CHAPITRE II

RAPPEL HISTORIQUES ET RESUME DES DIFFERENTES PHASES D'ETUDES

II.0 - LE POURQUOI DU A(DPCM)

Le choix délibéré de l'introduction d'adaptativité à plusieurs niveaux (stratégies de prédiction, prédicteurs, codage-canal...) a été guidé par les faits suivants:

- un savoir-faire était précédemment acquis au sein de notre laboratoire concernant l'utilisation d'algorithmes adaptatifs dans des problèmes de Traitement du Signal. Au sens "théorie de l'information" nous souhaitions obtenir en sortie des systèmes de codage étudiés dans le cadre de cette convention un signal d'erreurs de prédiction à entropie minimale et donc concevoir un système réalisant une extraction maximale des redondances spatiotemporelles du signal d'image $I(x,t)$.

- à l'inverse d'approches plus globales (notamment dans les algorithmes d'estimation et d'interprétation de mouvement dans une scène) utiles voire nécessaires dans des applications de Reconnaissance de formes et d'Analyse de scènes, l'approche locale est spécialement adaptée aux applications de codage pour lesquelles les critères visuels et le capteur de vision (l'oeil humain) qui détermineront in-fine la qualité d'images nécessaire, fonctionnent sur des contextes locaux.

- l'introduction d'adaptativité au sein des schémas de codage permet d'assurer une meilleure robustesse vis à vis de deux types de perturbations:

- > le bruit d'acquisition- et de quantification- source du signal
- > les erreurs de transmission potentiellement présentes sur le canal.

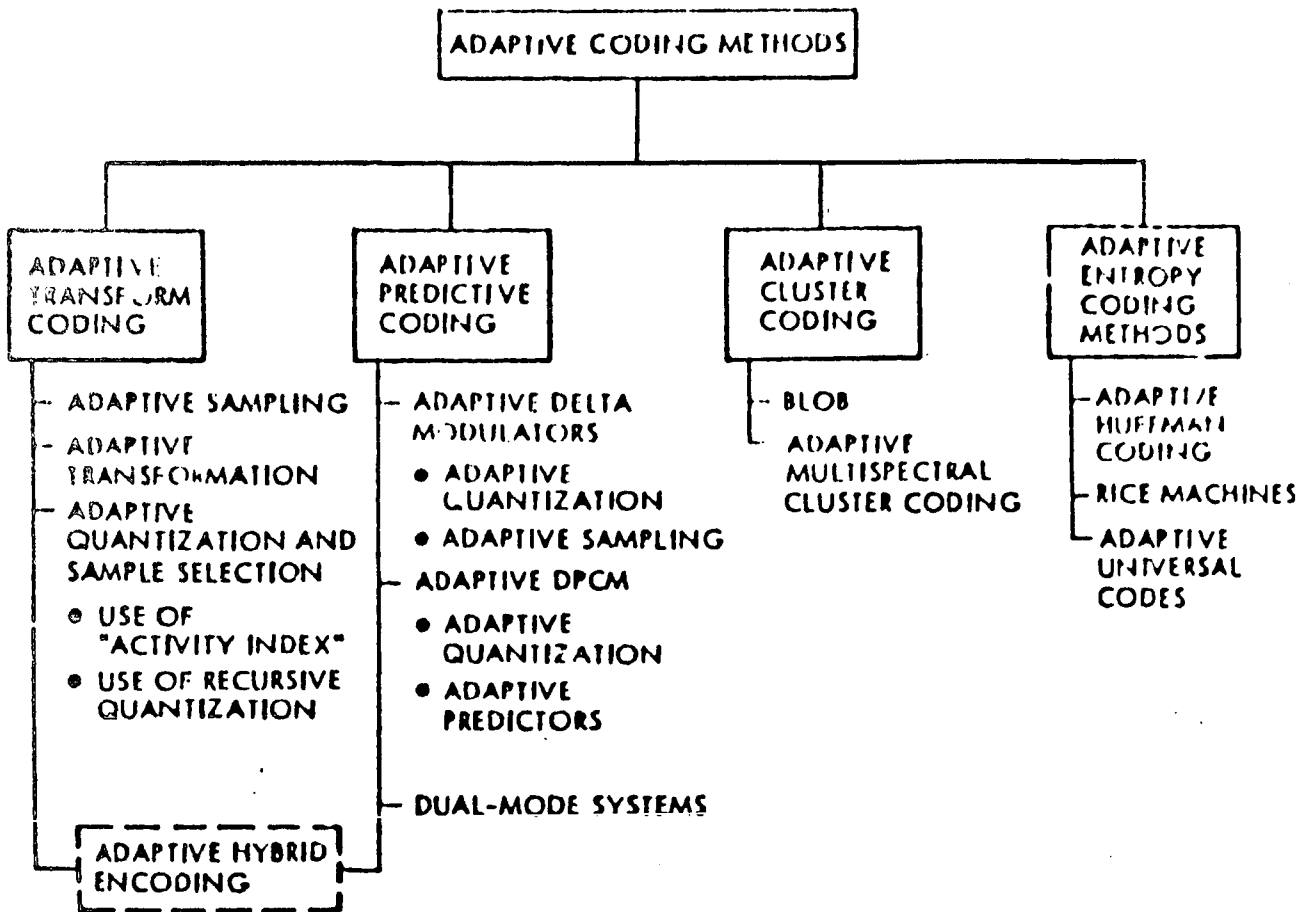


Figure 3

Introduction d'adaptativité dans les schémas de codage: extrait de [12]

Comme l'illustre la figure 3, l'introduction d'adaptativité au sein de schémas de codage peut être réalisée à plusieurs niveaux algorithmiques.

Pour notre part, nous l'avons introduit aux niveaux:

- du choix des modes de prédiction conduisant à des schémas de codage multimodes intra/interimage.

- du choix des prédicteurs: au sein d'un mode de prédiction, il est possible de définir une batterie de prédicteurs choisis de manière adaptative en chaque point selon le vecteur de caractérisation locale $C(x,t)$. Nous avons introduit cette possibilité au sein du mode de prédiction intraimage.

- du compactage variable des informations (codage par plages de longueur variable).

- du codage-canal qui peut être réalisé en utilisant les codes statistiques à longueur variable (type Huffman). Ce niveau d'adaptativité a été implicitement introduit à notre niveau par l'utilisation de l'entropie des signaux à transmettre (et donc la borne minimale de nombre de bits pour les codes statistiques) en tant que mesure de performances des schémas de codage définis.

Nous rappelons dans les paragraphes ci-dessous comment ont été introduits les différents niveaux d'adaptativité pour les modes de prédiction.

II.1 - HISTORIQUE N°1: SIMULATION D'UN CODEUR ADAPTATIF INTRAIMAGE

Le domaine de codage dans ce cadre reste purement intraimage. En reprenant les notations définies au chapitre I, ce schéma de codage adaptatif intraimage était caractérisé par:

- en chaque point ($x=(X,Y)$), une caractéristique $C(x)$ spatiale.

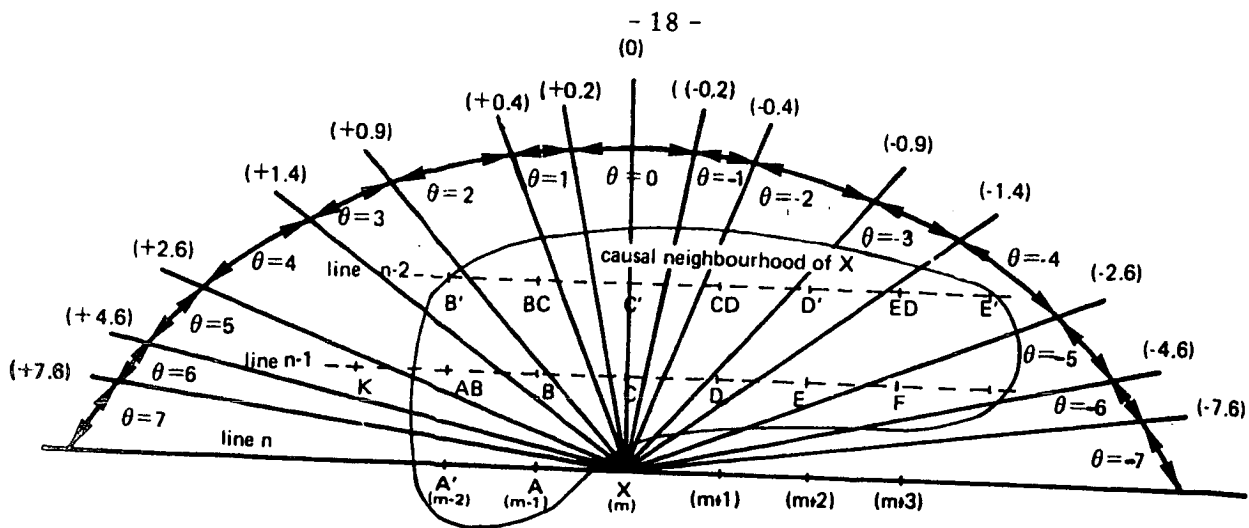


Figure 4: Voisinage causal d'un point d'image X

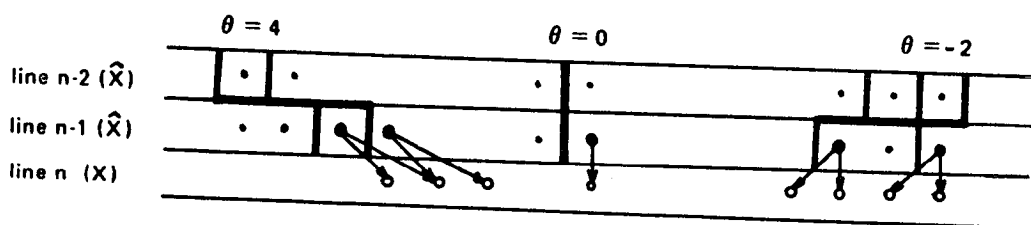


Figure 5 : Caractérisation locale prédictive

predictor \ θ	-4	-3	-2	-1	0	1	2	3	4	non edge pixels
algorithm B*	$\frac{D+E}{2}$	$\frac{3D+A}{4}$	$\frac{C+D}{2}$	$\frac{3C+D}{4}$	C	C	$\frac{B+C}{2}$	$\frac{A+C}{2}$	$\frac{A+B}{2}$	$\frac{A+C}{2}$
algorithm D	$\frac{A+D}{2}$ if $\theta < 0$				$\frac{A+C}{2}$ if $\theta \geq 0$ or non edge pixels					

Figure 6
Exemples de choix de prédicteurs fonction d'une caractérisation θ

- . la définition de voisinages $V_I^v(x)$ spatiaux (cf. figure 4).
- . le choix du prédicteur en fonction d'une caractérisation locale prédictive (cf. fig.5 et 6).

Définition de C(x)

La fonction locale de caractérisation C(x) spatiale du pixel x est définie par:

- en x, point contour ou non
- si point contour, estimation de l'orientation $\theta(x)$ locale du contour en ce point.

L'algorithme d'estimation de l'orientation des contours est réalisé par propagation d'un automate d'état associé à chaque contour dont:

- les mesures observées en entrée sont les gradients locaux horizontaux et verticaux
- les paramètres internes d'états de l'automate sont lissés par la propagation de celui-ci le long du contour et représentent ses caractéristiques: longueur, épaisseur,...
- l'estimée en sortie est la mesure de l'orientation lissée de ce contour.

Fonction de prédiction

La fonction de caractérisation C(x) est utilisée de manière prédictive pour définir la prédiction $\hat{I}(x)$ de I(x):

$$\hat{I}(x) = f(C(x), V_I^v(x))$$

avec dans le cas présent:

*C(x) = $\hat{\theta}(x)$, $\hat{\theta}$ orientation prédite au pixel x(X,Y) en fonction des estimées $\hat{\theta}(x'_i)$, x'_i étant les pixels de la ligne antérieure. La figure 5 illustre cette caractérisation locale prédictive.

* $V_I^j(x)$ est un voisinage causal spatial autour du pixel x à prédire

*deux exemples de fonction adaptative f sont donnés à la figure 6.

Résumés des enseignements de cette étude

Outre l'intérêt procuré au cours de cette étude par les phases de modélisation des contours d'une image par automate d'états et d'estimation de leur orientation, les performances de ce type de schéma de codage A-DPCM intraimage ont été évaluées en terme de débit de transmission résultant et de robustesse vis à vis d'erreurs de transmission.

Ainsi les résultats décrits dans [3], [8] montrent que pour une fréquence d'échantillonnage de 12Mhz en luminance, le débit moyen par pixel est de l'ordre de 3,5b/pixel pour une qualité parfaite en reconstruction. Cette réduction de débit peut atteindre 3b/pixel si une quantification adaptative est adjointe au schéma de prédiction adaptative.

Concernant les erreurs de transmission, ces schémas adaptatifs sont moins sensibles que les schémas de prédiction fixe ou que d'autres algorithmes connus comme Graham et Zschunke. Cependant leur comportement demeure critique si des prédicteurs bidimensionnels ne sont pas utilisés systématiquement par la fonction f de prédiction. Plusieurs variantes ont été ainsi testées et leurs sensibilités respectives comparées [2], [3].

II.2 - HISTORIQUE N°2: REALISATION HARDWARE D'UN CODEUR ADAPTATIF INTRAIMAGE

Sur la base des simulations précédentes, un algorithme de codage A-DPCM intraimage fut figé et traduit en terme d'architectures et de réalisation hardware. Les détails

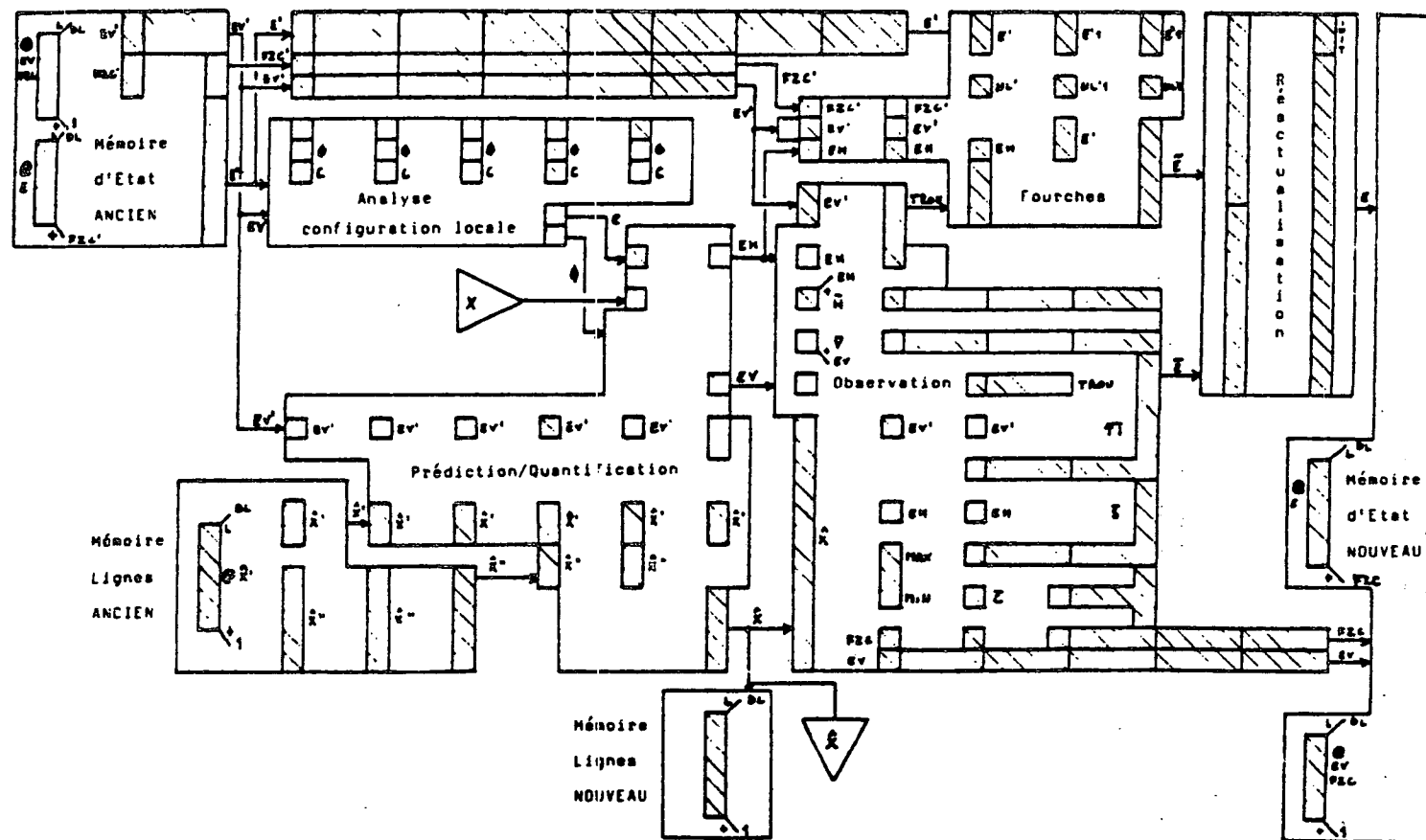


Figure 8 : Schéma temporel du flux de données extrait de [11]

de réalisation sont décrits dans [11]. Les figures 7 et 8 fournissent le schéma général d'architectures de ce codeur ainsi que l'enchaînement (pipe-line) temporel des étapes de traitement.

Du point de vue des technologies et horloges de synchronisation mises en jeu, deux blocs bien distincts sont à séparer au sein de cette réalisation hardware:

- le bloc "suivi de contours" qui réalise la gestion des automates d'états associés aux contours et fonctionne selon les horloges spécifiques à ces automates. La technologie employée fut les boîtiers TTL standard.

- la partie "Prédiction/Quantification", fonctionnant à l'horloge-pixel ($f_{smax} = 13,5 \text{ Mhz}$) la technologie utilisée pour ce faire fut des boîtiers ECL de la série Motorola 10000.

Les conclusions et conséquences de cette étude d'architecture furent de prouver la faisabilité en réalisation hardware de codeur "hautement" adaptatif au sens où nous l'avons défini au chapitre I et fonctionnant à la fréquence d'échantillonnage-pixel de 13,5Mhz. La réussite de cette entreprise n'était pas évidente à l'époque de la réalisation de cette machine. Depuis, la littérature abonde d'exemples d'autres réalisations de codeurs adaptatifs intraimages fonctionnant à de telles normes d'échantillonnage.

Le prototype hardware ne fut pas par la suite utilisé mais a permis, pour les études interimages et hybrides, la prise en compte, dès les simulations numériques, des contraintes Temps-réel dues à ces hautes fréquences d'échantillonnages et de la liaison existante entre complexité algorithmique et complexité de réalisation hardware.

II.3 - HISTORIQUE N°3: ETUDE MOUVEMENT

Les schémas de codage adaptatifs intraimages décrits précédemment visaient à définir des prédicteurs adaptés à toutes les configurations locales possibles du signal d'images, et principalement à combattre le mauvais comportement des prédicteurs fixes autour des non-stationnarités spatiales (contours). La même démarche a été prise dans le domaine spatiotemporel où les non-stationnarités spatiotemporelles correspondaient, cette fois, aux contours en mouvement.

L'étude "Mouvement" qui a été entreprise en vue de la définition de schémas de codage avec compensation de mouvement a donc comporté deux étapes:

- l'estimation du mouvement (et essentiellement de la composante perpendiculaire au contour) le long des contours par une modélisation et des schémas d'estimation analogues à ceux développés dans le cas intraimage pour l'estimation de l'orientation des contours.

- la détermination d'un champ complet de vecteurs de déplacement $d(x,t)$ estimés en chaque point d'image.

Nous rappelons brièvement quelques éléments sur ces deux phases d'études et leurs conclusions respectives.

II.3.a - Estimation de mouvement sur les contours

L'idée sous-jacente à cette étude est d'étendre la modélisation par automate d'état définie dans le cas bidimensionnel intraimage (cf. section II.1 et II.2) au cas tri-dimensionnel de l'espace (x,t) . Ainsi l'automate associé à chaque contour spatiotemporel est caractérisé par:

- *des mesures observées en entrée, constituées par les gradients locaux spatiaux et temporels (différences point à point intra et interimages)

- *des paramètres d'états filtrés récursivement lorsque l'automate parcourt spatialement le contour associé.

*une solution itérative par algorithme de gradient à ce problème de minimisation:

$$\hat{d}_{\text{point courant } x_i \text{ en mouvement}} = \hat{d}_{\text{point précédent } x_{i-1} \text{ en mouvement}}$$

$$\xi = \epsilon \text{DFD}(x_i, \hat{d}_{i-1}) \quad \begin{matrix} \text{gradient} \\ \text{temporel} \end{matrix} \quad \begin{matrix} V_I(x_i - \hat{d}_{i-1}, t-1) \\ \text{gradient} \\ \text{spatial} \end{matrix}$$

Nous avons souhaité améliorer les performances de convergence de ces algorithmes itératifs en introduisant sur tests de rupture de modèle (là où apparaissent des contours en mouvement et où ont pu être estimés des déplacements de contours) les estimations de mouvement calculées spécifiquement sur les contours par l'algorithme décrit en II.3.a.

II.3.c - Application au codage interimage avec compensation de mouvement

Le codage simple (monomode) interimage avec compensation de mouvement définit une fonction de prédiction $\hat{I}(x,t) = f(C(x,t), V_I^v(x,t))$ avec, dans ce cas:

$$C(x,t) = \hat{d}(x,t)$$

$$V_I^v = \text{voisinage centré autour de } x - \hat{d}(x,t) \text{ dans l'image } (t-1)$$

$$- f(C, V_I^v) = \hat{I}_{\text{interpolé sur } V_I^v(x - \hat{d}(x,t), t-1)}$$

Si aucune information sur les déplacements estimés n'est transmise, le schéma de prédiction doit fonctionner conformément à la figure 2 en utilisant une caractérisation locale $C(x,t)$ prédite. Cela signifie que $\hat{d}(x,t)$ dans les formules précédentes est à remplacer par $\hat{d}^v(x,t)$, mouvement prédit au point x à partir des estimées de mouvement des points causalement antérieurs.

Plusieurs variantes quant à la définition de cette caractéristique locale prédite ont été testées. La solution la plus simple demeure:

$$\tilde{d}(x,t) = \hat{d}(x-1,t)$$

avec $x = (X,Y)$

$x-1 = (X-1,Y)$ point correspondant à x à la ligne précédente dans l'image.

Le codeur interimage avec compensation de mouvement transmettra les erreurs quantifiées

$$\tilde{e}(x,t) = q(I(x,t) - \hat{I} \text{ interpolé sur } V_I^y(x-\tilde{d}(x,t), -t-1))$$

II.3.d - Conclusion de cette étude mouvement

Le gain apporté par des techniques de codage interimage a pu être mesuré [8], [9] en comparaison avec:

- les codeurs intraimages fixes ou adaptatifs
- les codeurs interimages simples c'est à dire sans compensation de mouvement.

Le gain est différent selon les zones d'image considérées: zones homogènes, contours, fixes ou en mouvement.

La figure 9 en donne une illustration en termes d'entropie et de probabilités d'erreurs.

De plus, le gain apporté par une estimation du mouvement des contours adjoint à l'algorithme "pel-récuratif" classique a lui aussi été mesuré en termes de réduction de la dynamique des erreurs de prédiction [7], [8]. Durant l'étude décrite au paragraphe II.4, l'algorithme d'estimation de mouvement utilisée dans le cadre de la compensation de mouvement reste la version complète: estimation de mouvement sur les contours et algorithme "pel-récuratif". Pour les études décrites au chapitre IV, le module estimation de mouvement des contours n'a pas été implanté.

intraimage				interimage		interimage avec compensation de mouvement	
$P_{\lambda} = \Pr(I(x,t) - \hat{I}(x,t) \leq \lambda)$							
Entropy E							
	A	C	$\frac{A+C}{2}$	FD	P_c	P_s	P_E
Entropy E	3.71	3.99	3.75	4.46	4.31	3.63	3.51
P_{60}	98.9%	96.5%	99.2%	92.8%	90.2%	96.9%	97.7%
P_{40}	96.8%	92.4%	97.2%	86.0%	83.7%	94.2%	95.5%
P_{20}	89.4%	83.2%	88.8%	72.7%	73.5%	88.0%	89.8%
P_{10}	78.6%	72.0%	75.8%	58.3%	63.8%	79.9%	82.0%

Ce choix a été opéré pour trois raisons principales:

- *la minimisation du temps CPU pour l'implantation de ces programmes sur le processeur ARTISTE (CCETT)

- *le souci de tendre vers une complexité moindre (réalisme pour une réalisation hardware) dans le cadre des phases finales d'évaluation des systèmes hybrides de codage intra/interimages.

- *enfin le fait que cette perte de performance au niveau de l'estimateur de mouvement doit être compensée par une meilleure adaptativité entre les différents modes de prédiction (intraimage, interimage sans et avec compensation de mouvement).

II.4 - HISTORIQUE N°4: SCHEMAS DE CODAGE MULTIMODES

L'optique de cette phase d'étude était de concaténer au sein d'un même schéma de codage l'ensemble des possibilités d'introduction d'adaptativité pour le choix des prédicteurs dans un domaine spatiotemporel.

II.4.1 - Choix des modes de prédiction

Dans un tel domaine spatiotemporel, deux modes "naturels" apparaissent immédiatement:

- le **mode intraimage** avec un choix adaptatif possible entre plusieurs prédicteurs spatiaux

- le **mode interimage**

Pour celui-ci, nous y avons distingué:

- le mode interimage sans compensation de mouvement (simple recopie temporelle de mémoire d'images)
- le mode interimage avec compensation de mouvement nécessitant un estimateur de mouvement.

*Nous n'évoquons systématiquement que les traitements interimages et non pas intertrames. Dans le cadre de cette étude, aucun traitement intertrame n'est effectué.

II.4.2 - Approche choisie pour le schéma de codage

Comme nous l'avons mentionné et argumenté au chapitre I, les options suivantes ont été prises:

***codage à débit variable:** pour chaque pixel, la même quantité d'informations n'est pas envoyée. Pour les points déclarés "fixes" ou "compensés en mouvement", aucune erreur de prédiction n'est envoyée. Seuls les points non compensés en mouvement se voient attribués une erreur de prédiction.

***approche locale prédictive:** En chaque pixel, une caractéristique locale $C(x,t)$ va être prédite en fonction des caractéristiques locales de pixels antérieurement reconstruits.

***trois modes de prédiction**

mode 1 : $\hat{I}(x,t) = \hat{I}(x,t-1)$ interimage

mode 2 : $\hat{I}(x,t) = I_{\text{interpolé dans } V_I^{\hat{v}}(x-\hat{d}(x,t),t-1)}$

interimage avec compensation de mouvement

mode 3 : $\hat{I}(x,t) = f(C(x) = \hat{\theta}, t)$ intraimage adaptatif

II.4.3 - Schéma général du codeur

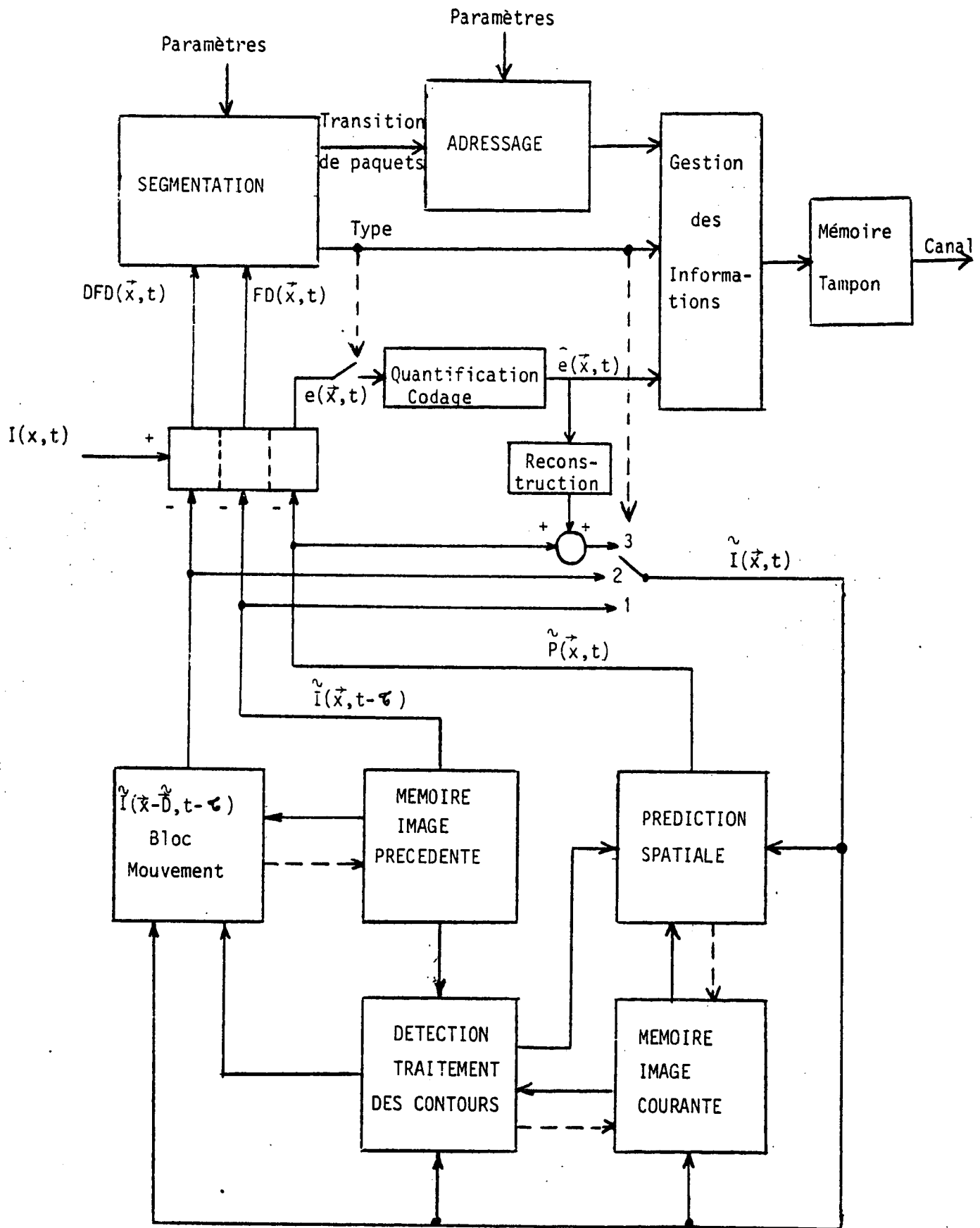
La figure 10 fournit le schéma général du codeur. Il est à noter que plusieurs variantes sur les blocs fonctionnels ont été testés [8]. Le chapitre III de ce rapport n'en fournira qu'un exemple. Notons ainsi que:

***pour le bloc "segmentation"**

Il s'agit de partitionner l'image entre les trois modes de prédiction 1, 2 et 3. Dans ce cadre deux méthodes principales ont été testées:

- une segmentation hiérarchique
- une segmentation avec hystérésis

C'est cette dernière qui a été retenue et que nous détaillons au chapitre III.



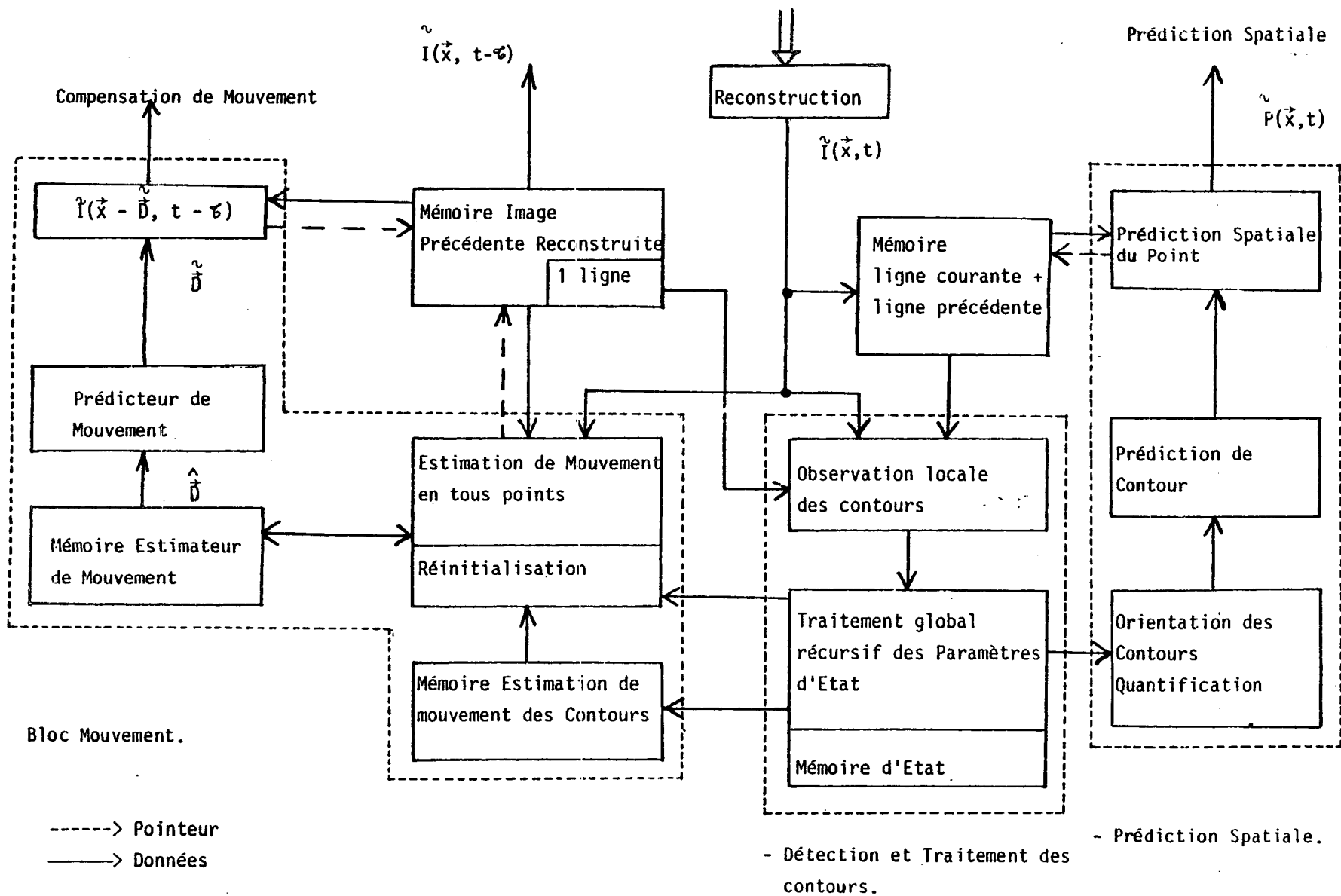


Figure 11: Schéma fonctionnel du bloc prédictif. extrait de [10]

***pour le bloc "estimateur de mouvement"**

La version complète (estimateur de mouvement sur les contours et algorithme d'estimation "pel-récuratif") a été employée dans la phase initiale d'études [10]. Seule la version simplifiée (estimation "pel-récuratif") a été implantée sur le processeur ARTISTE et sera détaillée au chapitre III.

***pour le bloc "prédiction spatiale"**

Trois variantes ont été implantées:

-prédiction fixe A

-prédiction fixe $\frac{A+C}{2}$

-prédiction adaptative: $\hat{I}(x) = f(\tilde{\theta}, V_I^y)$.

La caractéristique locale prédite $\tilde{\theta}$ étant l'information de présence ou absence de contours et orientation discrète (3 valeurs: $\tilde{\theta} < 0$, $\tilde{\theta} = 0$, $\tilde{\theta} > 0$).

Les résultats donnés ci-dessous utilise la variante adaptative. Le chapitre III utilisera la variante fixe $\frac{A+C}{2}$

II.4.4 - Type d'informations transmises

Le codage des diverses informations est effectué par plages (zones 1-D le long d'une ligne T.V) de point à point de prédiction identiques. Le codeur fournit ainsi trois types de plages:

plage de mode 1: points fixes

plage de mode 2: points compensés

plage de mode 3: "autres"

Pour chaque plage, l'information transmise est:

- n° du mode + adresse

Pour la plage de mode 3 seulement, est transmise une erreur de prédiction quantifiée.

	THRESHOLDS		ALGORITHM OF SEGMENTATION	NUMBER OF CLUSTERS/line	PERCENTAGE OF PIXELS		
	λ_1	λ_2			①	②	③
"JUDITH"	4	7	Hierarchical segmentation	19.5	88	9.5	2.5
			Segmentation with filtering	8.7	51.2	46.3	2.5
			Segmentation with filtering and sweeping	5.5	51	46	3
	3	5	Hierarchical segmentation	27	84	12	4
			Segmentation with filtering	14	50.5	45.5	4
			Segmentation with filtering and sweeping	8	50	45	5
"CASTLE"	7	12	Hierarchical segmentation	190	47	36	17
			Segmentation with filtering	103.6	13	70	17
			Segmentation with filtering and sweeping	52.5	11	68	21
	5	9	Hierarchical segmentation	229	37	40	23
			Segmentation with filtering	144	13	64	23
			Segmentation with filtering and sweeping	68	10	61	29

Figure 12
Résultats sur les séquences "Judith" et "Chateau"
des méthodes de segmentation

II.4.5 - Principaux résultats et enseignements de cette phase d'études

La figure 12 fournit un exemple de résultats produits sur les séquences "Judy" (visiophone) et "castle" (images TV, maquette).

La définition d'un tel schéma de codage a visé à atteindre deux objectifs:

1. utiliser au mieux les configurations locales spatiotemporelles du signal $I(x,t)$ pour le prédire de manière la plus efficace en fonction des différents domaines de prédiction possibles. Cet objectif a été atteint: le pourcentage de pixels partitionnés dans le mode 3 (seul mode où une erreur de prédiction est transmise) a été réduit au maximum sans perte de qualité de l'image en reconstruction.

2. réduire le nombre de plages par ligne afin de réduire le volume du débit binaire de transmission consacré aux adresses. Les différentes améliorations (détaillées en [10]) apportées au bloc de segmentation permettent une réduction conséquente de ce volume.

II.5 - CONCLUSION PARTIELLE

Nous avons jugé utile d'effectuer ces divers rappels historiques et méthodologiques afin d'explicitier la version figée algorithmique qui fait l'objet du chapitre suivant et dont les résultats joints illustrent le comportement sur séquences normalisées de T.V.

La conclusion générale que nous pouvons donner à l'issue de ces différentes phases d'études est de peser les différents éléments dans le compromis: complexité algorithmique - réduction de débit réalisée.

*le gain apporté par la compensation de mouvement n'est plus à démontrer. De nombreux exemples [6], [7] dans la littérature ont montré et quantifié la réduction en termes

de débit d'erreurs de prédiction (pour notre cas, de l'ordre de 1 bit d'entropie [9] sur séquence T.V. complexe). Toutefois la difficulté reste la réalisation effective de l'estimateur de mouvement associé au compensateur de mouvement. Peu de réalisations hardware ont vu le jour. Nous avons, notamment pour cette raison, pris une version simplifiée dans les simulations suivantes.

*En jouant de manière adaptative sur le choix du prédicteur dans le domaine spatiotemporel le plus large possible (seule restriction à celui-ci, des bornes maximales dans le schéma d'estimation de mouvement sont imposées afin d'éviter les divergences rédhibitoires de l'estimateur; cependant des pixels situés dans ces zones sont très rarement affectés au mode 2 utilisant l'estimateur de mouvement), nous pensons avoir cerné le plus près possible les corrélations spatiotemporelles locales exploitables pour une prédiction optimale. Dans ce contexte, seul un changement de domaine de travail (domaine fréquentiel par transformation orthogonale) peut apporter d'autres atouts dans la recherche d'un prédicteur optimal.

*Comme cela a été effectué dans le cadre des schémas de codage A-DPCM monomode intraimage, nous avons souhaité observer le comportement "dynamique" d'un algorithme de codage A-DPCM multimodes en présence:

- de différentes séquences d'entrée à caractéristiques (rapport S/N, taux de mouvement...) diverses

- d'erreurs de transmission sur le canal entre le codeur et le décodeur.

Tel est le sujet du chapitre III qui fournit les résultats de simulation concernant ces questions.

CHAPITRE III

EXPERIMENTATIONS COMPLEMENTAIRES

III.1 - SIMULATIONS SUR SEQUENCES "COST"

III.1.1 - Détails algorithmiques sur le schéma de codage utilisé

Nous allons préciser successivement les différents blocs algorithmiques de la figure décrivant le schéma général du codeur multimodes utilisé dans le cadre de ces simulations.

a. Estimateur de mouvement

L'algorithme "pel-recursif" décrit dans [5] a été utilisé

$$\hat{D}^i = \hat{D}^{i-1} - \text{sign}(\text{DFD}(x_i, \hat{D}^{i-1})) \times \text{sign}(\nabla I((x_i - \hat{D}^{i-1}), t-1))$$

avec $\epsilon = 1/8$

-1 pour $x < 0$

$\text{sign}(x) = 0$ pour $x = 0$

1 pour $x > 0$

La récursivité i a lieu dans le **sens ligne** du balayage. L'initialisation par le vecteur de vitesse nulle est effectuée au début de chaque trame et au début de chaque ligne (ce qui revient à l'hypothèse de bords de l'image fixes).

Un tel schéma d'estimation fournit des vecteurs de mouvement estimés avec une partie fractionnaire.

Le calcul des "DFD" nécessite donc une procédure d'interpolation. Rappelons que:

$$DFD(x,d) = I(x,t) - \tilde{I}_{\text{interpolé dans } v_I^{\sim}(x-d,t-1)}$$

$\tilde{I}_{\text{interpolé}}$ étant calculé sur l'image des intensités reconstruites stockée dans la mémoire d'image.

Afin d'éviter des divergences trop élevées, des bornes (+/- 7 pels/trame) contraignent la formule d'estimation précédente. Si un dépassement en estimation apparaît, la réinitialisation au vecteur de déplacement nul est effectuée. Ces bornes ont été fixées en majorant les vitesses maximales observées sur la séquence "Voiture".

b. Schéma d'interpolation

Nous reprenons le schéma d'interpolation défini en ([5] p.659) basé sur une triangulation (avec somme pondérée pour le calcul de la luminance estimée) entre les trois plus proches voisins au sens de la distance euclidienne.

Ainsi pour le calcul des "intensités déplacées", $I(x-d,t-1)$ les poids dans le calcul de l'interpolation seront fonction directe des parties fractionnaires dans les directions horizontale et verticale du vecteur de déplacement.

$$I = I_D + D'_1(I_C - I_D) + D'_2(I_B - I_D).$$

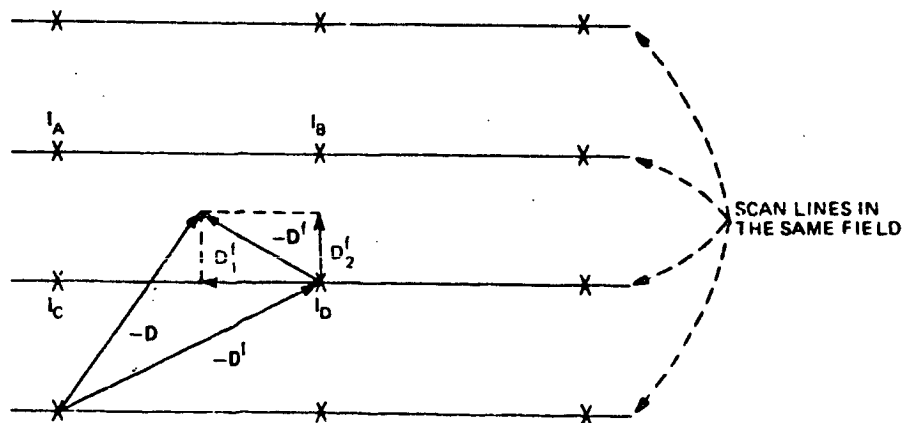


Figure 13: Schéma d'interpolation. extrait de [5]

c. Caractéristique locale prédite

En chaque pixel x , le choix du mode de prédiction va être orienté (voir § segmentation) sur la base des informations locales suivantes:

***mouvement local prédit pour le pixel x**

$$\tilde{d}(x,t) = \hat{d}(x-1,t)$$

$(x-1)$ étant le pixel correspondant à x sur la ligne précédente.

***différence interimage locale**

$$DF(x,t) = I(x,t) - \tilde{I}(x,t-1)$$

\tilde{I} étant une intensité reconstruite.

***différence interimage déplacée**

$$DFD(x,\tilde{d}) = I(x,t) - \tilde{I}_{\text{interpolé dans } V_I^V(x-\tilde{d},t-1)}$$

d. Principes généraux pour l'algorithme de segmentation

L'algorithme de segmentation retenu est un algorithme partitionnant l'image en trois modes:

mode 1 : mode fixe

mode 2 : mode compensé en mouvement

mode 3 : mode spatial (du fait du choix d'un prédicteur spatial pour celui-ci)

Il réalise:

.une segmentation avec hystérésis

.un nettoyage (des "petites" plages segmentées) sans dégradation objective

.un nettoyage avec dégradation mais non perceptible visuellement

Nous précisons ces différents points ci-dessous.

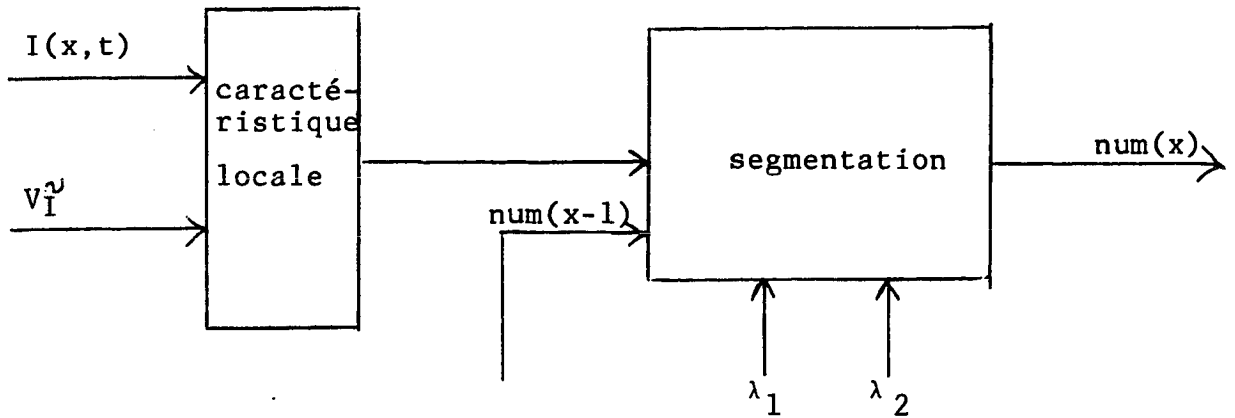


Figure 14: Segmentation avec hystérésis

e. Segmentation avec hystérésis

Les figures 14 et 15 présentent l'algorithme de segmentation retenu. La référence [11] fournit une autre présentation de celui-ci ainsi que les différentes variantes testées.

.la variable $\text{num}(x)$ désigne le numéro de mode affecté au pixel

.l'hystérésis est introduit par la prise en compte de l'information $\text{num}(x-1)$

.les paramètres λ_1 et λ_2 ont été choisis sur histogrammes et observations visuelles des images après reconstruction. Dans le cas présent,

$$\lambda_1 = 7, \lambda_2 = 12$$

Pour résoudre les effets de bords, il est toujours supposé l'hypothèse de bords fixes soit $\text{num}(0) = 1$.

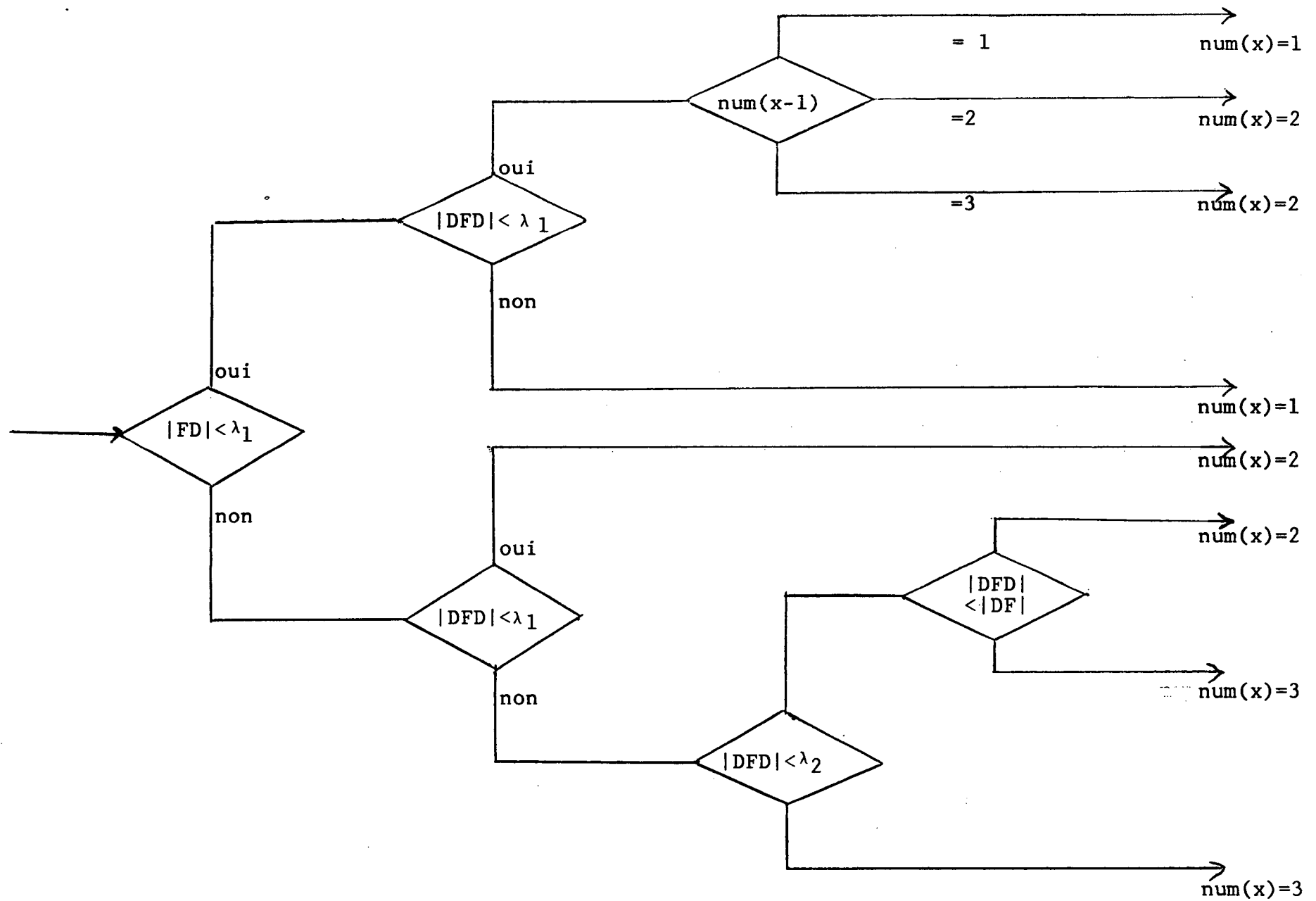


Figure 15
Segmentation avec hysteresis décrite par arbre de décision séquentiel

f. Procédure de nettoyage des plages segmentées

L'idée sous-jacente à l'introduction de telles procédures est de minimiser le volume d'adresses de plages à transmettre en pontant les plages de tailles trop petites (longueur=1 ou 2 pixels). Les pontages peuvent être réalisés:

.sans dégradation objective: dans ce cas, il y a substitution d'informations (d'adresse + numéro de mode) par une erreur de prédiction transmise (mode 3). Ce nettoyage vers le mode 3 sera ainsi réalisé pour les configurations de plages suivantes:

33333XX3333

33333X3333

333XYYYY

YYYX3333

333333

3333333

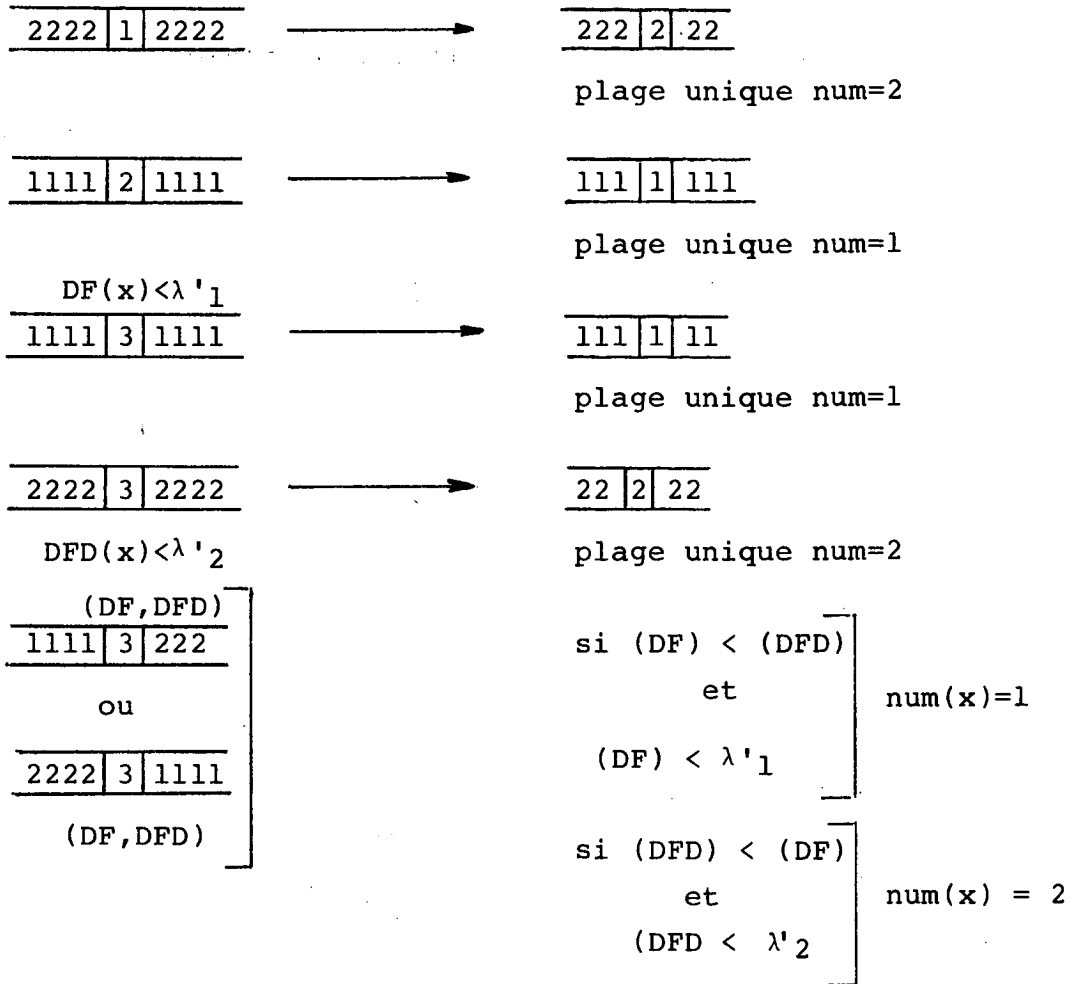
3333YYYY

YYY33333

(avec retour 1 passe)

.sans dégradation visuelle

Une prédiction sous-optimale est choisie dans ce cas; cependant des précautions ont été prises pour que les "dégradations objectives" injectées ainsi ne soient pas critiques visuellement. La procédure de pontage retenue est la suivante:



Dans les simulations réalisées, λ'_1 et λ'_2 ont pris les valeurs 10 et 12.

g. Choix des prédicteurs

Les trois modes de prédiction définis par le processus de segmentation se voient associés les prédicteurs suivants:

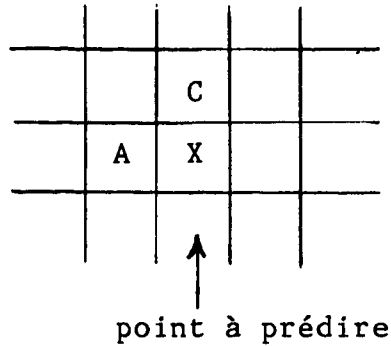
mode 1 : $\hat{I}(x, t) = \tilde{I}(x, t-1)$

mode 2 : $\hat{I}(x, t) = \tilde{I}_{\text{interpolé}} \text{ sur } v_I^{\check{v}}(x-\check{d}, t-1)$

mode 3 : $\hat{I}(x,t)$ = prédiction spatiale

$$= \frac{A+C}{2} \quad \text{sauf} \quad \begin{aligned} &= C \text{ (1er pixel de chaque ligne)} \\ &= A \text{ (1ère ligne de chaque trame)} \end{aligned}$$

avec les notations pour les pixels :



h. Quantification

Le mode 3 (points non compensés) est le seul mode de prédiction où sont transmises après quantification l'erreur de prédiction

$$\hat{e}(x,t) = Q(I(x,t) - \hat{I}(x,t))$$

Aucune recherche d'optimisation d'une caractéristique de quantification n'a été menée dans le cadre de cette étude interimage. Nous avons repris les valeurs de seuils et niveaux de reconstruction optimisés pour un quantificateur intraimage à 15 niveaux | |, | |.

q_i -75 -55 -39 -25 -14 -6 -2 2 6 14 25 37 50 60 86

s_i -64 -46 -31 -19 -10 -4 0 4 10 19 30 43 57 75

Caractéristique de quantification

Un mot de code est affecté pour indiquer la fin de plage en mode 3. Le mode 3 utilise ainsi 16 mots de code pour transmettre les erreurs de prédiction soit:

$$n_{\text{bits}}/\text{err} = 4 \text{ bits}$$

III.1.2 - Résultats des simulations

a. Données utilisées

Les données de séquences d'images utilisées furent les trois séquences normalisées au niveau du groupe COST 211 bis, échantillonnées à 13,5Mhz en luminance et possédant chacune 90 images successives. La figure fournit l'illustration en contenu de scène et type de mouvement sur ces séquences. Leurs caractéristiques sont les suivantes:

***séquence Voiture:** cette séquence est la plus critique en termes de taux et d'amplitude du mouvement. Elle possède des structures spatiotemporelles périodiques (grille en mouvement) fortement critiques du point de vue estimation de mouvement. Les mouvements sont de nature frontaux comme transversaux (déplacement de la voiture).

***séquence Musi et Girl:** ces deux séquences ont un taux de mouvement moins critique. Le mouvement sur la séquence Girl est plus équirépartie alors que la séquence Musi présente un "CUT" (changement de scènes brutal à l'image n°75) intéressant à étudier du point de vue du comportement du codeur à débit variable.

La fréquence d'échantillonnage pixel conduit à travailler sur des matrices-trame de points de dimension (675 pixels x 268 lignes). La luminance est codée sur 8 bits et comprise dans l'intervalle $[16, 235]$.

b. Simulations réalisées

Les simulations réalisées sur ces séquences ont visé à:

.fournir une indication sur le fonctionnement dynamique de ce schéma multimodes avec les observations de variabilité du débit binaire résultant en sortie ainsi que l'évolution des répartitions des pixels entre les trois modes de prédiction.

.fournir une évaluation en terme de débit (dans un but comparatif avec d'autres schémas de codage). Nous tenons à préciser que le train réel binaire de transmission n'est pas produit en sortie de ces simulations en l'absence de définition des algorithmes de codage-canal pour l'adressage et de la gestion du buffer-tampon. Il s'agit donc ici d'une évaluation de débit calculée avec les hypothèses suivantes:

1. nombre de bits par erreur de prédiction transmise:

4 bits (cf. III.1.1.4)

2. nombre de bits pour l'adressage (relatif monodimensionnel): adressage + numéro de type: (cf. ||1|)

5 bits

d'où le calcul:

débit binaire par pixel =

$$4 \times \text{Prob}(\text{num}(x)=3) + \frac{\text{nb plages par ligne}}{675} \times 5$$

c. Résultats et commentaires

La figure 17 illustre le caractère fortement local du choix des modes de prédiction.

Les différentes étapes de l'algorithme de segmentation sur la séquence Voiture et Girl et leurs conséquences sont illustrées respectivement en figure 18 et 19 en termes de répartition par modes de prédiction et nombre de plages par ligne.

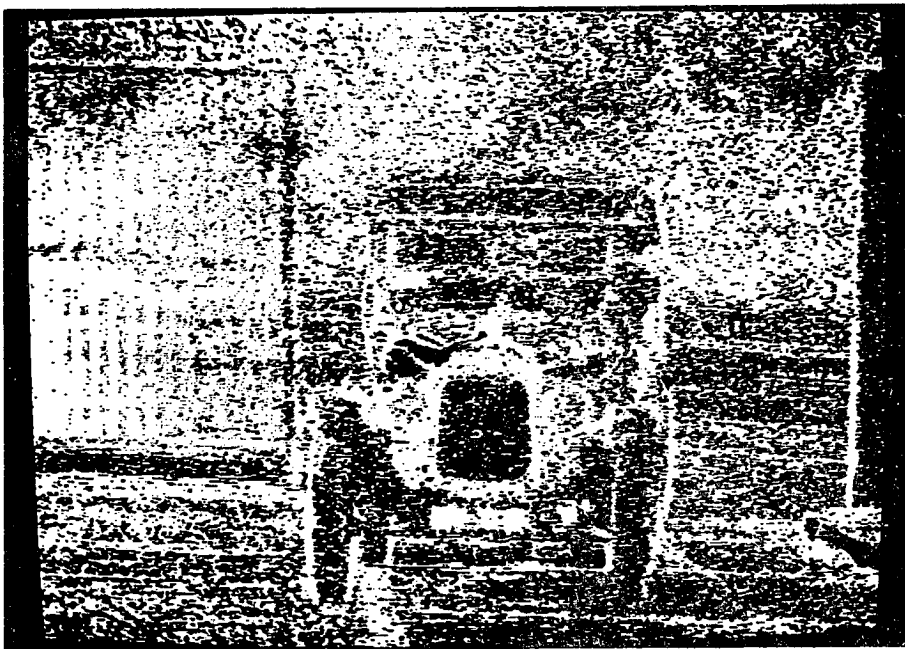


Figure 17 : Images "Modes de prédiction" sur Voiture

n°image	n°trame	Segmentation avec hysteresis					
		Segmentation hiérarchique		sans dégradation objective		sans dégradation visuelle	
		nb paquets/ ligne	T1, T2, T3 en %	nb paquets/ ligne	T1, T2, T3 en %	nb paquets/ ligne	T1, T2, T3 en %
2	1	108	86 7 7	64	55 32 12	49	56 33 11
	2	112	86 7 7	65	53 34 13	51	53 35 12
3	1	124	83 8 9	69	50 34 16	55	51 35 14
	2	129	82 9 9	70	48 36 16	56	49 36 15
4	1	139	80 10 10	74	45 38 17	60	45 39 16
	2	145	77 11 12	73	43 37 20	60	44 37 19
5	1	150	74 12 14	72	39 38 23	60	40 39 21
		153	71 13 16	72	37 38 25	59	38 38 24

Figure 19
Résultats de l'algorithme de segmentation sur quelques trames (GIRL)

n°image	n°trame	Segmentation hiérarchique		Segmentation avec hysteresis			
				sans dégradation objective		sans dégradation visuelle	
		nb paquets/ ligne	T1, T2, T3 en %	nb paquets/ ligne	T1, T2, T3 en %	nb paquets ligne	T1, T2, T3 en %
71	1	323	40 30 30	124	16 36 48	104	16 37 47
	2	325	41 30 29	125	16 36 48	105	16 37 47
72	1	319	42 29 29	124	16 36 48	104	16 37 47
	2	321	42 29 29	122	16 36 48	103	16 37 47
73	1	318	42 29 29	121	16 36 48	103	16 37 47
	2	319	42 29 29	122	16 36 48	103	15 37 48
74	1	319	42 29 29	124	16 37 47	104	15 38 47
	2	320	42 29 29	123	16 37 47	104	16 38 46
75	1	320	43 29 28	125	16 37 47	105	15 39 46
	2	320	43 29 28	124	16 37 47	104	15 39 46

Figure 18
Résultats de l'algorithme de segmentation sur quelques trames (VOIT.)

La figure 20 fournit pour les séquences Musi (séquence image 1 à 50), Musi-cut (séquence image 71 à 80), Girl (séquence image 1 à 50) et Voiture (séquence image 1 à 10 et 71 à 79) deux types de résultats:

- *la répartition (en %) entre les trois modes de prédiction,
- *l'évaluation du débit de transmission.

Sur ces dernières figures, les courbes en pointillé concernent les débits obtenus en utilisant un algorithme de segmentation hiérarchique simple (1ère variante introduite dans [11]). Les courbes en trait plein sont les évaluations de débit en utilisant l'algorithme de segmentation défini dans le paragraphe précédent.

Ce débit total (courbe supérieure) est décomposé en débit d'adresses (partie supérieure entre les deux courbes) et débit d'erreurs de prédiction (dû au mode 3: courbe inférieure).

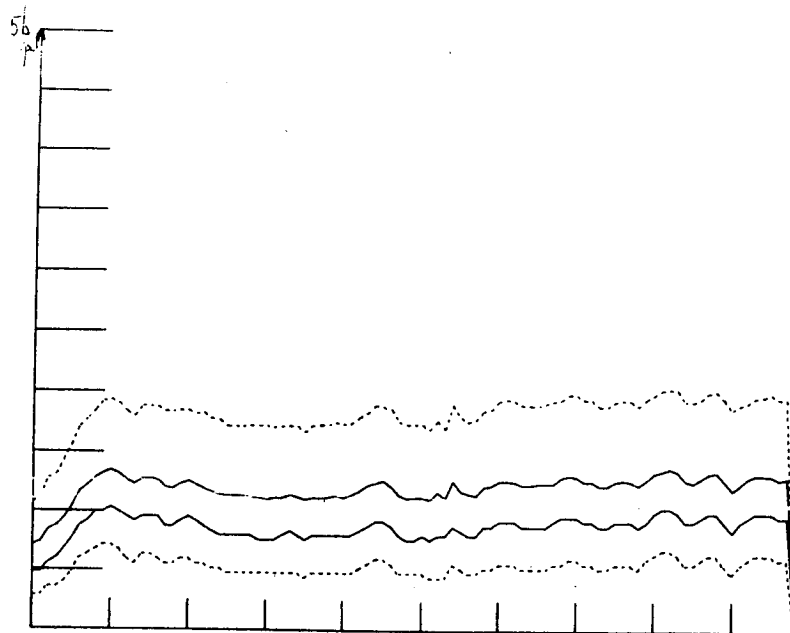
De ces résultats, peuvent être tirés les enseignements suivants:

1. l'algorithme de segmentation retenu réalise un gain en débit total appréciable (par rapport à la méthode de segmentation hiérarchique simple) en réduisant fortement le débit dû aux adresses au prix d'une faible hausse du débit d'erreurs de prédiction,

2. à l'inverse de [5], la part dans le débit total du débit d'adresses n'est pas majoritaire. Ce résultat est principalement dû aux procédures de segmentation et nettoyage employées,

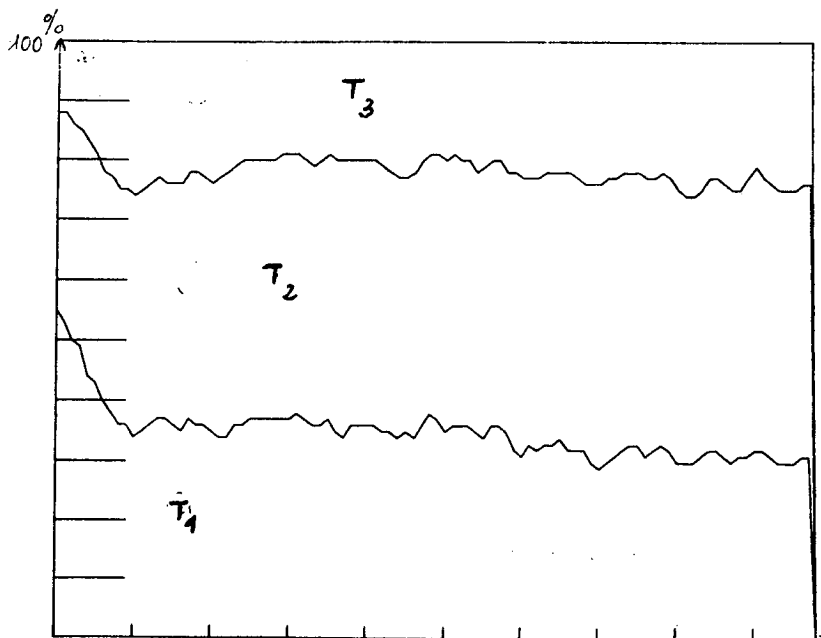
3. le débit variable résultant du fonctionnement en boucle ouverte (sans buffer-tampon) de ce codeur est stable pour un contenu de scène identique. Le résultat est infirmé dans les trois situations suivantes:

.initialisation de séquences (10 premières trames de la séquence Girl): le taux de mouvement détecté semble faible en début de cette séquence ce qui crée un débit plus faible au départ et une répartition-pixel privilégiant



GIRL
(100 trames)
Evaluation du
débit (b/pixel)
en utilisant
la segmentation
hiérarchique et
la segmentation
avec hysteresis

a

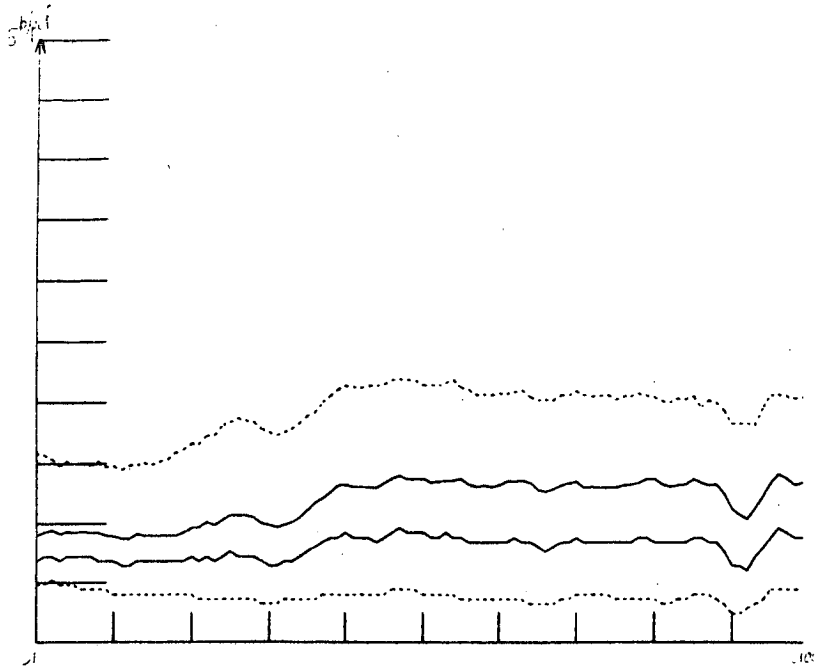


b

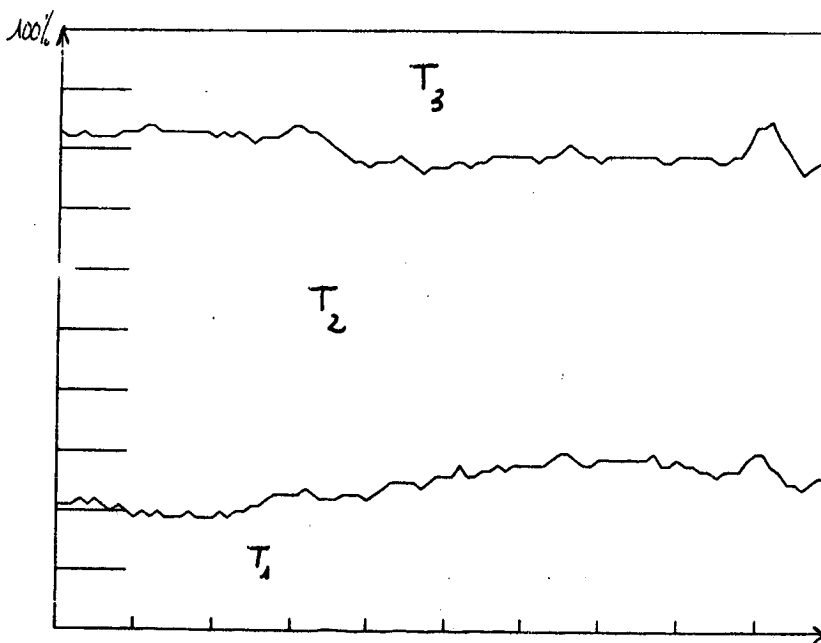
Figure 20: Répartition des pixels entre les 3 modèles de prédiction

MUSI
(100 trames)

Evaluation du
débit (b/pixel)
en utilisant la
segmentation hié-
rarchique et la
segmentation avec
hysteresis



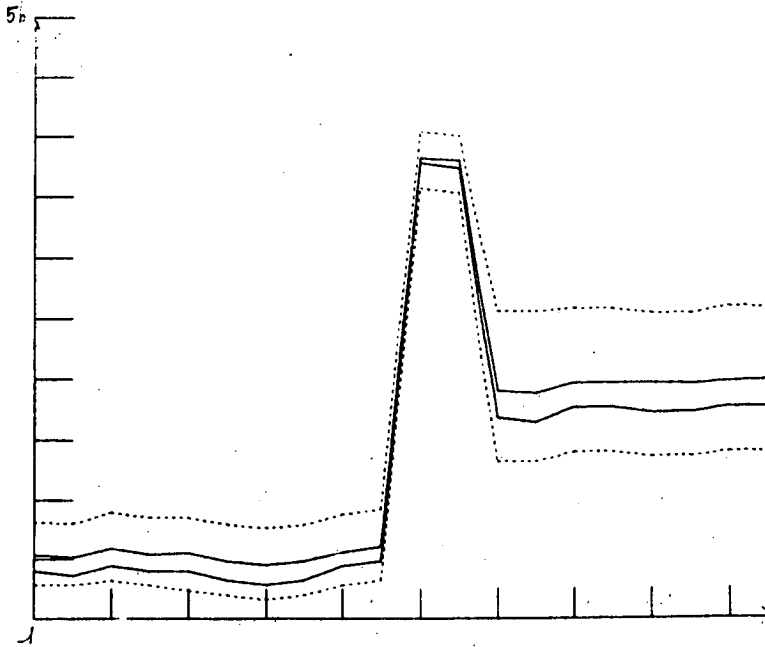
c



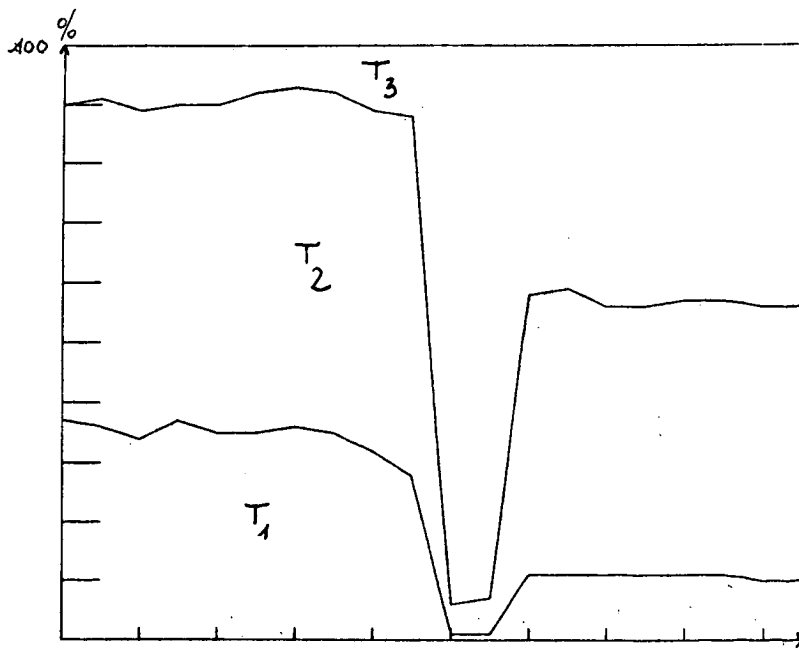
d

MUSI-CUT
(10 trames)

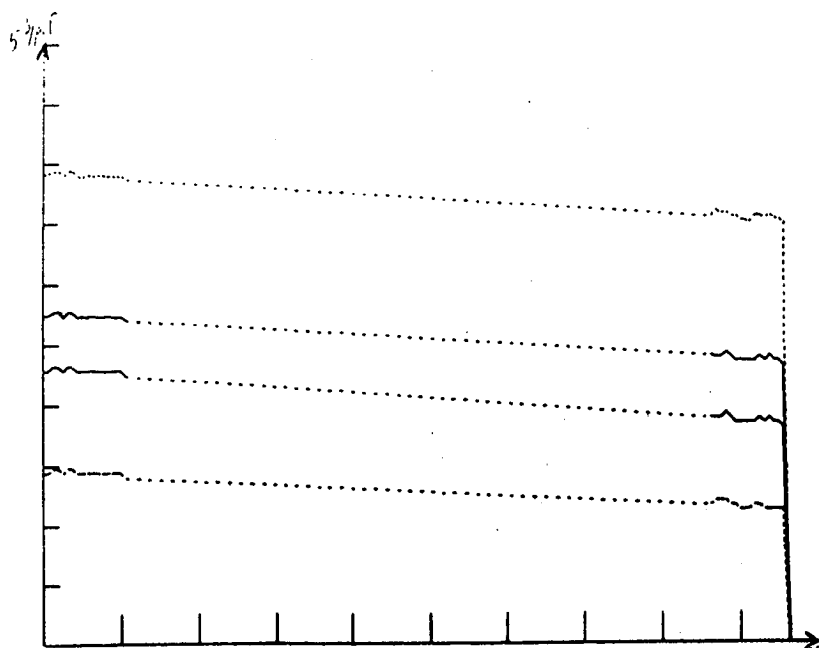
Evaluation du
débit(b/pixel)
en utilisant la
segmentation hié-
rarchique et la
segmentation avec
hysteresis



e

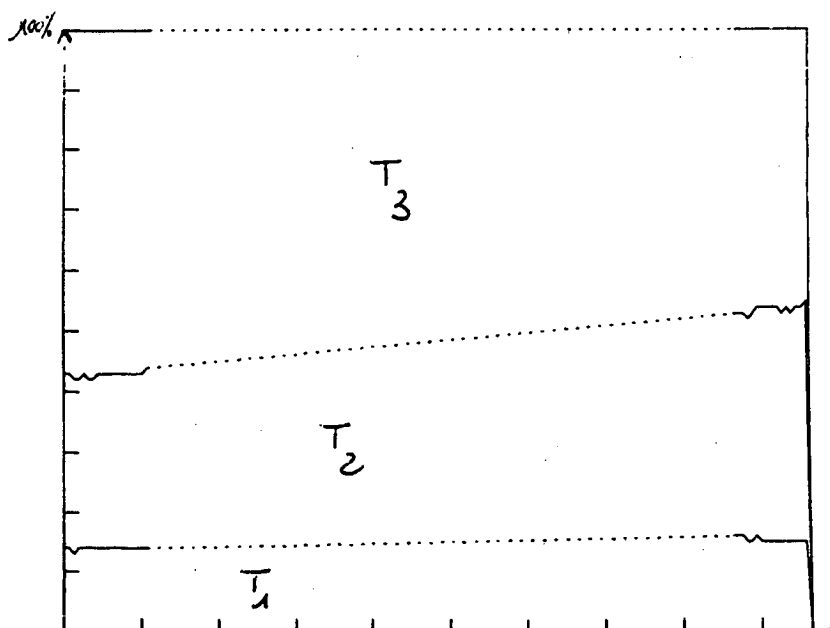


f



VOITURE
(160 trames; les
simulations sur
les trames en
pointillé sont
extrapolées)

8



h

les modes de prédiction temporelle. Ce phénomène n'est pas observé pour les autres séquences,

.traversée d'un "cut" (changement brusque de contenu de scène): la transition image n°75 image n°76 de la séquence Musi nous fournit un tel cut. On y observe, de part et d'autre du "cut", deux régimes stables à débit total différent (0,5b/pixel -> 1,5b/pixel) et à répartition par mode différent. Cela rejoint l'expérimentation menée dans [14] de création d'un "cut" synthétique par concaténation de deux séquences différentes. Ce résultat montre aussi la dépendance du débit variable d'un tel schéma adaptatif et son fonctionnement interne (répartition des pixels par mode de prédiction) au contenu spatiotemporel (beaucoup de gradients et de mouvements) de la scène traitée. La figure 21 ci-dessous donne quelques résultats globaux comparatifs approximés en moyenne sur les séquences traitées.

Répartition par modes de prédiction				Evaluation de débits		
	mode 1	mode 2	mode 3		débit moyen	débit max.atteint
Voit	15%	35%	50%	Voit	2,5b/pix	2,77b/pix
Girl	35%	30%	25%	Girl	1,25b/pix	1,38b/pix
Musi	25%	55%	20%	Musi	1,25b/pix	1,41b/pix

Figure 21: Comparaison globale

.changement de séquences: la figure 21 montre le comportement différent du codeur face à trois types de séquences.

La séquence "Voiture" reste la séquence la plus critique: seulement 50% des pixels ne nécessitent aucune transmission d'erreur de prédiction. Du fait de l'important taux de mouvement présent dans cette séquence, le mode 2 avec compensation de mouvement est le mode temporel le plus souvent utilisé.

La séquence "Musi" est peu critique et autorise des réductions de débit importantes. Ses mouvements semblent fortement prédictibles (55% pour le mode 2).

La séquence "Girl" présente des performances intermédiaires: le débit résultant est faible (1,25bit/pixel) mais la compensation de mouvement est dans ce cas moins efficace globalement que le simple rafraichissement temporel.

4.Comparaison avec d'autres approches: les résultats obtenus semblent cohérents avec ceux obtenus sur les mêmes séquences par d'autres approches. Ainsi dans [14], nous pouvons relevé les mêmes conclusions concernant la séquence Voiture (débit avoisinant 2,3b/pixel) et le caractère moins critique des séquences Girl et Musi. La traversée du "cut" dans Musi donne également des résultats similaires. Il y a toutefois des différences quant à la répartition par modes de prédiction: l'algorithme de segmentation et les critères de décision sont différents; le caractère non prédictif de l'estimation de mouvement utilisé dans la compensation de mouvement privilégie également les modes temporels.

5.Sans perte en qualité d'images après reconstruction, l'utilisation de plusieurs modes de prédiction et leur

sélection adaptative en fonction des configurations locales du signal permettent d'obtenir d'importantes plages de pixels (80% des pixels pour Musi) ou aucune innovation spatiotemporelle n'est à transmettre au décodeur. Ce résultat correspond au souci affiché en préalable à cette étude (cf. Introduction du présent rapport) d'extraire au maximum les redondances spatiotemporelles présentes au sein de la séquence.

6.Enfin, il est à noter que l'approche adoptée qui est fortement orientée "Traitement adaptatif du signal d'images", ne donne que peu d'informations interprétables en terme d'Analyse de scènes et d'identification/paramétrisation de formes. Ainsi la partition de l'image entre les trois modes ne fournit en aucun cas une segmentation cohérente de la scène en zones fixes, zones en mouvement homogène,... Les valeurs prises par l'estimateur de mouvement sont elles-aussi peu interprétables en terme de vecteurs de déplacement réels.

III.2 - ROBUSTESSE D'ALGORITHMES ADAPTATIFS MULTIMODES INTRA/INTERIMAGES AUX ERREURS DE TRANSMISSION

(étude réalisée par J.P. de la Tribonnière)

III.2.1 - Introduction

Ces simulations ont fait l'objet d'exposés et publications dont les résumés sont donnés en Annexe 7, 8 et 9.

Il nous a paru intéressant de compléter le lot de simulations décrites en IV.1 par une observation du comportement d'un schéma de codage tel que celui détaillée précédemment. La phase d'observation des propagations temporelles d'erreurs de transmission à travers un schéma complet codeur-décodeur se devait d'être complétée par quelques expérimentations sur les méthodes potentielles de correction (suppression ou atténuation) des effets des erreurs de transmission au sein de la séquence reconstruite. Ces méthodes ont été regardées tout d'abord sans contrainte de débit additionnel et également avec la contrainte de 5% (du débit total) représentant le débit binaire alloué pour réaliser ces corrections-atténuations.

La méthodologie pour mener à bien cette étude a été en partie celle adoptée en [2] pour l'étude des effets des erreurs de transmission dans le cas intraimage.

III.2.2. - Hypothèses expérimentales

Nous supposons les hypothèses suivantes concernant le type d'erreurs introduites et le train binaire transmis:

- *les erreurs se présentent de manière isolée et non en paquets

- *le taux d'erreurs choisi pour ces expérimentations est compris entre $2.8 \cdot 10^{-6}$, $4 \cdot 10^{-4}$ ce qui correspond au minimum à une erreur injectée par image et au maximum 144 erreurs par image. Ces taux sont supérieurs aux taux

classiquement rencontrés dans les canaux de transmission; cependant, afin d'observer l'évolution dynamique des perturbations au sein d'une petite séquence (20 images par exemple), il convenait d'injecter un nombre conséquent d'erreurs de transmission. Pour des raisons de rapidité d'exécution des simulations, bien souvent celles-ci ont été menées sur un quart d'image ce qui correspond pour le taux maximal d'erreurs à 18 erreurs injectées par trame (6 par modes de prédiction et également injection aléatoire).

*les synchronisations Image, Trame, Ligne sont supposées fortement protégées. Le codeur est supposé posséder la faculté d'autodétection d'erreurs de synchronisation du train binaire; ainsi pour une ligne d'image le train binaire a une cohérence interne en fonction des modes de prédictions et des longueurs de plages obtenues; en cas de détection d'erreurs dans la cohérence de ce train, il y a rejet du train ligne et remplacement par des techniques de dissimulation que nous présentons ci-dessous.

L'ensemble des simulations a été réalisé sur la séquence Voiture.

III.2.3 - Techniques de dissimulation intraimage et interimage

Les techniques de dissimulation classiquement utilisées travaillent dans le domaine intraimage [2] mais doivent respecter les contraintes de causalité dans le cas où des prédictors bidimensionnels spatiaux sont utilisés. Dans le domaine interimage, une dissimulation temporelle peut être envisagée et méritera d'être testée avec toutefois la contrainte similaire de causalité temporelle.

Nous avons testé deux types de dissimulation:

.la dissimulation de la ligne courante déclarée erronée par la ligne précédente de la même image (dissimulation intraimage)

INTRAFRAME AND INTERFRAME DISSIMILATIONS :

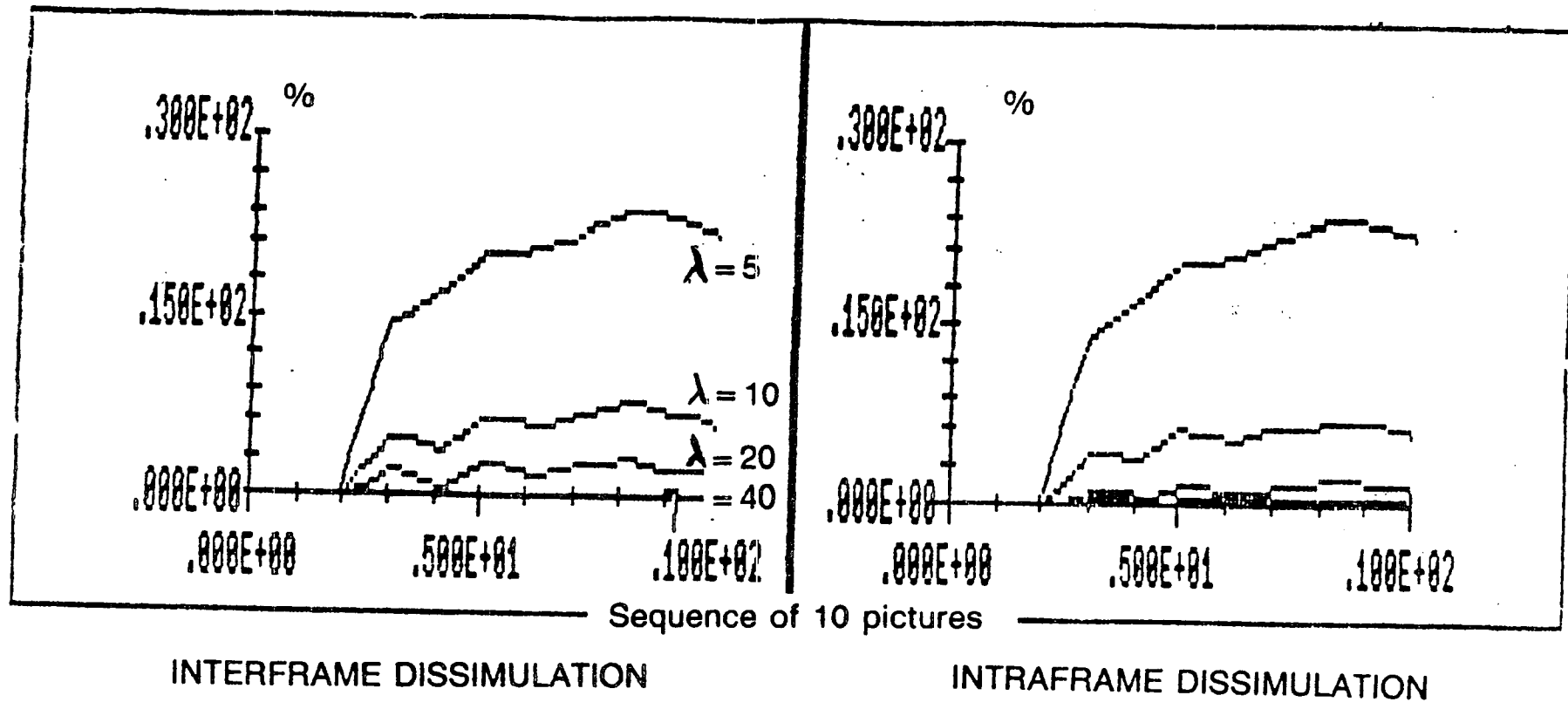


Figure 22

Simulations comparatives de dissimulation spatiale et temporelle sur 10 images

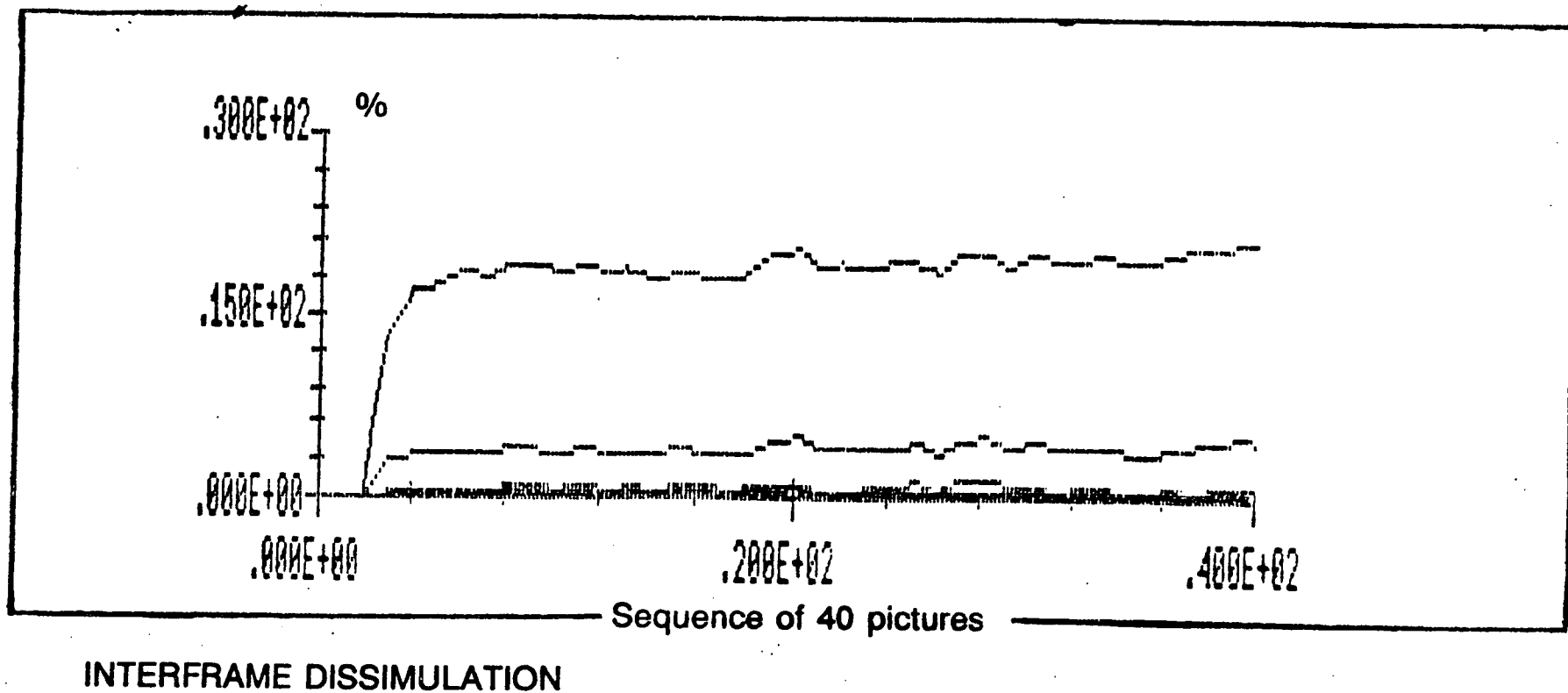


Figure 23
Dissimulation interimage testée sur 40 images de la séquence VOITURE

Toutefois, ces configurations sont **locales** et ne peuvent donc rendre les performances de telle ou telle dissimulation meilleures **de manière globale** sur l'ensemble de la ligne dissimulée.

Il est cependant observé des perturbations fortement visibles dans les zones-contours en mouvement (introduction de décrochages dûs au mouvement), ex: grille dans la séquence Voiture).

Les résultats joints en Annexe IX ainsi que ceux donnés par les figures 22 et 23 fournissent essentiellement trois types d'enseignements:

- *statistiquement, les techniques de dissimulation intraimage et interimage ont des performances identiques en terme d'erreurs introduites après reconstruction (ex: comparaison des deux courbes de la figure 22),

- *très rapidement (env. 5 images après l'introduction de dissimulation) un régime asymptotique est atteint,

- *du point de vue perception visuelle, les différences entre les images reconstruites sans et avec erreur ont été seuillées avec différentes valeurs de seuils ($\epsilon = 5, 10, 20$ et 40). La valeur 10 , fournit ainsi environ 6% de pixels d'"erronés". La dégradation reste donc visible.

La figure 23 illustre le comportement de la technique de dissimulation interimage sur une longue séquence (40 images de Voiture). Les lignes dissimulés sont au nombre de 10 lignes par image et leurs positions tirées aléatoirement sans répétition.

III.2.4 - Techniques de correction

L'observation brute de la propagation immédiate (dans l'image où les erreurs se sont introduites) des erreurs

IMAGE NO 3																			
CLASSES CODEUR POINT 474, LIGNE 152																			
2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	1	1	1
2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	2	1	1
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	1	1	1	3	3	1	1	1	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	1	1
2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1
CLASSES DECODEUR																			
2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	1	1	1
2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	2	1	1
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	1	1	1	3	3	1	1	1	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	1	1	1
2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1
CODEUR																			
92	84	86	82	83	132	131	117	100	124	187	140	56	35	19	30	63	176	17	
90	85	79	94	143	146	136	123	123	176	187	177	166	151	171	186	193	170	16	
98	85	76	99	158	150	129	122	150	177	184	174	176	177	172	173	169	167	16	
91	86	79	139	162	131	132	133	155	180	163	165	156	162	169	162	161	163	16	
92	75	102	157	145	136	123	132	163	160	158	163	160	153	144	156	164	161	16	
86	94	135	160	150	149	124	142	192	162	162	163	166	165	148	165	163	156	16	
82	86	135	153	137	118	127	159	163	155	157	148	132	152	155	152	151	155	15	
97	128	163	185	147	126	163	163	138	140	123	110	135	129	140	132	135	151	14	
98	99	119	113	116	187	225	192	202	208	190	200	204	200	189	176	168	156	13	
88	91	99	100	69	53	64	103	138	134	137	129	91	70	54	113	226	213	21	
DECODEUR																			
93	90	87	83	94	139	132	113	102	128	194	150	64	38	16	24	62	171	17	
95	90	82	96	145	148	133	130	130	179	192	185	174	156	172	184	189	166	15	
101	89	79	101	116	80	95	106	143	175	185	179	182	183	175	173	167	164	16	
93	89	82	141	92	61	80	99	135	169	163	165	159	162	169	162	161	163	16	
96	78	105	160	112	94	76	39	137	139	145	157	160	153	190	156	164	161	16	
87	96	137	162	135	115	81	99	153	133	141	155	159	162	170	177	168	158	16	
84	88	137	155	131	98	95	122	163	155	157	142	125	152	155	152	154	158	15	
100	131	171	188	145	115	142	134	123	133	120	106	129	126	138	131	136	153	14	
102	102	122	116	116	181	211	170	193	195	182	194	198	200	189	176	163	156	13	
89	93	101	102	70	50	55	87	121	119	125	120	84	67	53	112	226	213	21	
DIFFERENCE DECODEUR-CODEUR SEUIL= 4																			
0	6	0	0	0	0	0	0	0	0	7	10	8	0	0	-6	-6	-5	0	
5	5	0	0	0	0	0	0	0	0	5	3	9	5	0	0	0	0	0	
0	0	0	0	-142	-70	-34	-16	-7	0	0	5	6	6	0	0	0	0	0	
0	0	0	0	-70	-70	-52	-34	-20	-11	-5	0	0	0	0	0	0	0	0	
0	0	0	0	-33	-52	-52	-43	-31	-21	-13	-6	0	0	46	0	0	0	0	
0	0	0	0	-15	-34	-43	-43	-37	-22	-21	-13	-7	0	22	11	5	0	0	
0	0	0	0	-3	-20	-32	-37	0	0	0	-6	-7	0	0	0	0	0	0	
0	0	0	0	0	-11	-21	-29	-15	-7	0	0	-6	0	0	0	0	0	0	
0	0	0	0	0	-6	-14	-22	-19	-13	-3	-6	-6	0	0	0	0	0	0	
0	0	0	0	0	0	-7	-16	-17	-15	-12	-9	-7	0	0	0	0	0	0	

Figure 24

Injection d'une erreur (amplitude 150)
sur l'erreur de prédiction (classe ③)
Observation locale des perturbations introduites

IMAGE NO 3																			
CLASSES CODEUR										POINT 261/LIGNE 114									
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
CLASSES DECODEUR																			
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
CODEUR																			
111	91	37	90	97	92	89	90	37	108	94	80	85	63	54	60	81	97		
189	190	188	164	155	179	174	167	169	152	125	104	69	64	61	66	98	72		
194	183	185	182	187	181	174	176	182	186	158	102	71	75	64	63	56	39	1	
189	192	202	190	173	182	183	193	195	179	154	103	73	72	54	56	58	54	16	
137	139	131	121	94	99	143	154	164	139	107	80	62	61	63	73	63	33	12	
229	234	232	226	207	206	199	197	184	141	102	89	52	52	60	62	57	53		
174	165	184	180	179	198	196	198	189	140	96	32	56	55	64	62	53	47	12	
225	224	218	216	215	211	217	213	199	130	88	83	55	60	73	62	48	30	16	
215	217	211	215	206	206	213	207	178	115	94	66	54	53	55	64	42	42	18	
181	177	170	171	174	190	194	194	172	118	67	64	54	56	66	59	40	43	16	
DECODEUR																			
111	91	37	90	97	92	89	90	37	108	94	80	85	63	54	60	81	97		
189	190	188	164	155	179	174	167	169	152	125	104	69	64	61	66	98	72		
194	183	185	182	183	181	174	176	182	186	158	102	71	75	64	63	56	39	10	
189	192	202	190	173	182	183	193	195	179	154	103	73	72	54	56	58	54	16	
137	139	131	121	94	99	143	154	164	139	107	80	62	61	63	73	63	33	12	
229	234	232	226	207	206	199	197	184	141	102	89	52	52	60	62	57	53		
174	165	184	180	179	198	196	198	189	140	96	32	56	55	64	62	53	47	12	
225	224	218	216	215	211	217	213	199	130	88	83	55	60	73	62	48	30	16	
215	217	211	215	206	206	213	207	178	115	94	66	54	53	55	64	42	42	18	
181	177	170	171	174	190	194	194	172	118	67	64	54	56	66	59	40	43	16	
DIFFERENCE DECODEUR-CODEUR SEUIL= 4																			
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	11	30	11	5	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	17	14	10	5	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	8	11	10	9	5	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	3	10	0	-5	0	6	9	23	1	
0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	15	1	
0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 25

Injection d'une erreur sur le numéro de classe (②→①)
Observation locale des perturbations introduites

IMAGE NO 3																					
CLASSES	CODEUR	POINT	227	LIGNE	136																
1	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	2	2	2	3	
2	2	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
3	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2	3	3	3	3	3	
1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	2	2	3	3	3	3	
2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	2	2	3	3	3	3	
3	3	3	3	3	3	3	3	2	2	2	1	1	1	1	3	3	3	3	3	3	
3	3	3	1	1	1	2	2	2	2	1	1	1	1	3	3	2	2	2	2	2	
CLASSES		DECODEUR																			
1	3	3	3	3	3	3	3	3	2	2	2	1	2	2	2	2	2	2	2	3	
2	2	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
3	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2	3	3	3	3	3	
1	2	2	1	1	1	1	1	1	1	1	1	1	1	2	2	3	3	3	3	3	
2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	2	2	3	3	3	3	
3	3	3	3	3	3	3	3	2	2	2	1	1	1	1	3	3	3	3	3	3	
3	3	3	1	1	1	2	2	2	2	1	1	1	1	3	3	2	2	2	2	2	
CODEUR																					
143	145	153	158	143	140	162	155	149	159	155	153	154	156	157	156	155	141	15			
153	159	159	170	168	170	171	167	163	162	159	160	159	161	153	153	162	162	16			
166	175	169	167	170	175	167	167	177	163	165	163	163	164	165	167	175	164	16			
153	161	163	151	162	171	163	171	163	170	169	170	183	171	170	132	181	186	18			
134	133	142	144	147	134	146	133	136	159	150	162	156	146	154	150	149	142	13			
168	154	151	149	141	139	143	165	151	150	151	153	150	147	156	157	164	173	17			
164	166	169	166	164	163	157	153	154	154	156	157	168	157	161	170	161	163	18			
143	149	150	150	175	167	168	170	165	170	172	167	154	150	157	157	145	143	15			
71	55	63	81	103	96	77	94	106	101	102	111	126	136	132	138	136	142	15			
180	167	165	172	169	173	168	164	170	157	161	162	158	161	152	146	143	144	14			
DECODEUR																					
143	145	153	153	143	140	162	155	149	159	155	153	154	156	157	156	155	141	15			
158	159	159	170	168	170	171	167	163	162	159	160	159	161	158	153	162	162	16			
165	175	169	167	151	155	164	162	152	156	163	163	163	166	170	167	175	164	16			
158	161	163	151	153	156	154	154	152	156	161	164	177	169	171	193	181	186	17			
134	133	142	144	142	124	137	125	124	145	139	153	156	146	154	150	149	142	13			
163	154	151	149	141	139	144	159	151	151	150	153	151	152	159	168	164	173	16			
164	166	169	166	164	163	157	153	154	154	156	157	163	157	161	170	161	163	15			
148	149	150	150	175	167	168	170	166	170	172	167	154	150	157	157	145	148	15			
71	55	63	81	103	96	77	94	106	101	102	111	126	136	132	138	136	142	15			
180	167	165	172	169	173	168	164	170	157	161	162	158	161	152	146	143	144	14			
DIFFERENCE		DECODEUR-CODEUR SEUIL= 4																			
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	-19	-21	0	-5	-25	-12	0	-5	-5	0	5	0	0	0	0	0	
0	0	0	0	0	-9	-15	-9	-7	-16	-14	-8	-3	-6	0	0	0	0	0	0	0	
0	0	0	0	0	-5	-10	-9	-8	-12	-13	-11	-9	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	-6	0	0	0	0	0	5	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Figure 26

Injection d'une erreur sur le numéro de classe (①->②)
 Observation locale des perturbations introduites

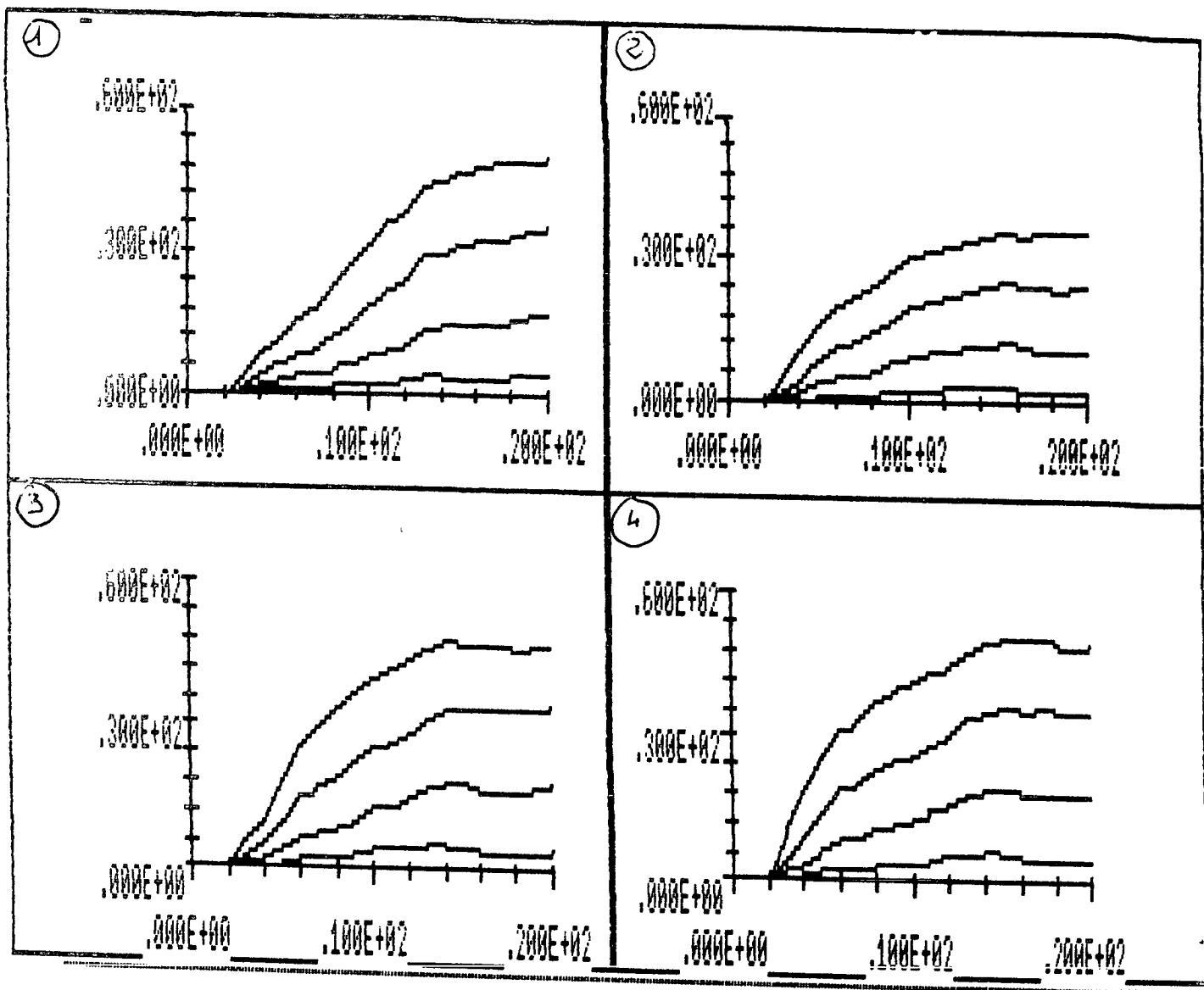


Figure 27

Courbes des pourcentages de points tels que la différence en valeur absolue du signal après le codeur et après le décodeur et injection d'erreurs, supérieure au seuil s ($s = 5, 10, 20, 40$)

Disposition des 4 graphiques

- ① classe 1 densité d'erreurs $D = 1, 3 \cdot 10^{-4}$
- ② classe 2 densité d'erreurs $D = 1, 3 \cdot 10^{-4}$
- ③ classe 3 densité d'erreurs $D = 1, 3 \cdot 10^{-4}$
- ④ somme des classes 1, 2, 3 $D = 3, 9 \cdot 10^{-4}$

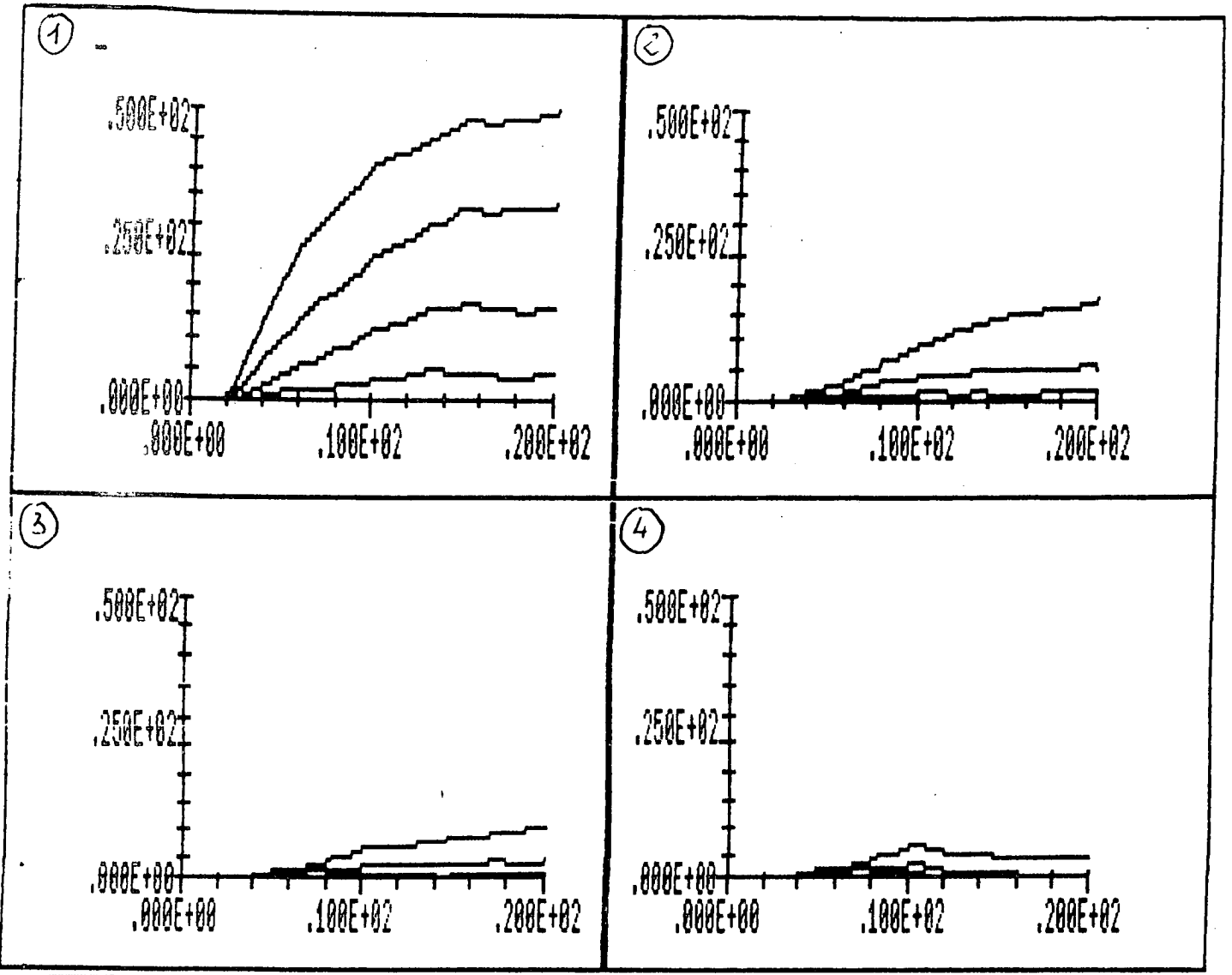


Figure 28

- ① correction 4 lignes "justes", erreurs sur les images 3-20
- ② correction 1 vecteur D "juste" pour 20 points
- ③ correction 1 vecteur D "juste" pour 10 points
- ④ correction vecteur D "juste" pour 10 points, erreurs sur les images 3-10

Disposition

TRANSMISSION OF DISPLACEMENT VECTORS :

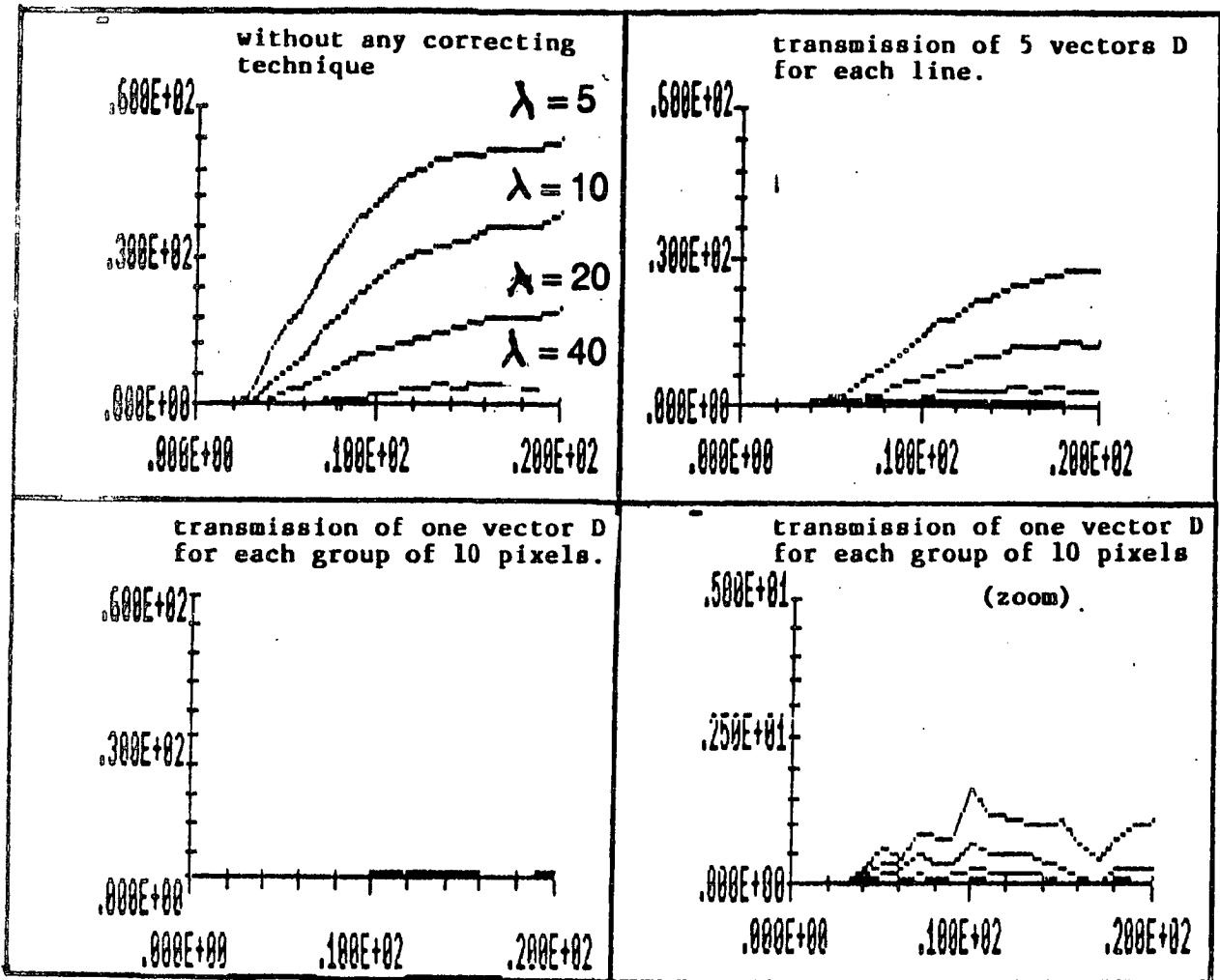


Figure 29

Simulations comparatives de diverses techniques de correction en présence d'un flux continu d'erreurs sur 20 images

de transmission a été effectuée sélectivement selon la nature du mode (1, 2 ou 3) de prédiction où sont injectées les erreurs. Les figures 24, 25 et 26 en fournissent quelques exemples locaux.

Les causes de propagation temporelle au sein de la séquence sont très diverses et sont analysées en détail dans l'annexe IX. Nous pouvons remarquer néanmoins que le caractère hautement adaptatif du codeur réalise un brassage très rapide de l'ensemble de ces causes de propagation. La figure 27 confirme ce résultat en obtenant des performances asymptotiques similaires quelque soit le niveau d'injection des erreurs.

Deux types de techniques de correction ont été testées:
une correction par l'envoi d'informations PCM de luminances. La contrainte de 5% de débit additionnel pour un tel envoi revient à transmettre 10 lignes PCM par images,

*une correction par l'envoi de vecteurs de déplacements: nous avons supposé l'envoi de vecteurs de déplacement (), chaque composante étant codée sur 4 bits

La contrainte de 5% de débit additionnel pour un tel envoi revient alors à transmettre un vecteur D tous les 60 pixels. Les simulations réalisées sur 1/4 d'image (388 pixels par 135 lignes pour une trame) signifiaient l'envoi de 5 vecteurs de déplacements par ligne d'image.

Des essais comparatifs ont été mené pour mesurer le pouvoir correcteur de ces envois de vecteurs de déplacement (1 vecteur "juste" sur les 10, 20 et 60 pixels) et sur des séquences de 10 à 20 images avec flot continu et interrompu d'erreurs (cf. Fig.28 et 29).

*Ces contraintes de débits ont été calculées sur la base du débit maximal obtenu pour coder la séquence Voiture soit 2,77b/pixel (cf. Fig.21).

III.2.5 - Conclusion partielle

Au vu de ces expérimentations, il apparaît nettement que:

*tant pour les techniques de dissimulation que pour les techniques de correction l'effet perturbateur créé par la propagation temporelle d'erreurs de transmission ne peut être annihilé complètement; tout au plus, il est observé un comportement asymptotique pour lequel les différences dans les images reconstruites restent visibles ($\lambda = 5$ ou 10 : nombre de points erronés importants).

*dans le cas d'un flot rare ou interrompu d'erreurs de transmission le système adaptatif "corrigé" se restabilise sur l'image reconstruite correcte (asymptotes nulles pour $\lambda = 40, 20, 10$; seul demeure un bruit granulaire d'erreurs pour le seuil $\lambda = 5$ ($\approx 5\%$)).

*sur l'ensemble des techniques de correction testées, celle qui apparaît être la plus efficace est la transmission additionnelle de quelques vecteurs de déplacements. En effet dans de tels algorithmes comprenant notamment un mode interimage avec compensation de mouvement, la propagation temporelle des erreurs de transmission semble en grande partie se faire dès la phase d'estimation-prédiction des vecteurs de déplacements sans resynchronisation possible sur des vecteurs de déplacement transmis correctement, il y a effet cumulatif et durable des erreurs en reconstruction suite à une divergence entre le champ de vecteurs de déplacement estimé au codeur et celui estimé au décodeur.

Une telle conclusion va dans le sens des algorithmes transmettant sur le canal des vecteurs de déplacement par blocs. L'étude en présence d'erreurs de transmission régulière pour chaque bloc de pixels des vecteurs de déplacement reste néanmoins à faire; ces informations nécessitent sans doute une protection particulièrement coûteuse, le décodeur dans de tels algorithmes, ne pouvant réestimer lui-même les vecteurs de déplacement.

CONCLUSIONS ET PERSPECTIVES

Ce rapport a eu pour objet d'effectuer le bilan synthétique de l'ensemble de nos études menées dans le cadre des schémas de codage MICD-A. Nous avons souhaité inclure la présentation des derniers résultats obtenus sur les séquences COST au sein de l'ensemble des études de recherche menées dans ce cadre.

A l'issue de ces diverses recherches expérimentales, il est possible de conclure sur les points suivants:

- *pour des séquences d'images de Télévision à contenu de scène complexe et mouvement important, la réduction de débit qui peut être atteinte semble avoisiner un facteur de 4 (luminance sur 8 bits \rightarrow luminance sur 1.5 à 2.5 bit/pixel). L'introduction de modes de prédiction temporelle permet ainsi de réduire d'environ 1 bit à 1,5 bit/pixel le débit de transmission nécessaire à qualité visuelle d'images égale par rapport à l'utilisation du seul mode de prédiction intraimage.

- *les modes de prédiction temporelle (avec et sans compensation de mouvement) sont choisis dans le cas des schémas de codage multimode étudiés pour 75 à 80% des pixels dans le cas de séquences à taux de mouvement moyen et pour 45 à 50% des pixels dans le cas de séquences à taux de mouvement important. **Les corrélations temporelles sont donc, dans tous les cas, importantes et méritent d'être utilisées et extraites pour la conception de schémas de codage efficaces.**

*a contrario, il demeure néanmoins toujours des zones non-stationnaires spatiotemporellement (situées sur les contours d'occlusion en mouvement) non aisément prédictibles tant dans le domaine interimage qu'intraimage.

Cette conclusion est systématiquement vérifiable lorsque l'on observe l'image d'erreurs de prédiction ou l'image des points où une information d'innovation reste à transmettre: sur de telles images, les structures "contours" réapparaissent naturellement.

Cette remarque conduit à la conclusion que, peut-être, un plafond est atteint dans le domaine d'extraction maximale des redondances spatiotemporelles présentes dans une séquence; ce plafond ne pourrait encore être dépassé qu'au prix d'importantes complexités algorithmiques (introduction d'estimateurs de mouvement spécifiques sur ces zones critiques, suivi temporel de ces structures,...) qui restent à comparer par rapport à d'autres approches tout aussi compétitives: ce constat a été ressenti de manière large suite aux diverses présentations du Picture-Coding-Symposium 86 de Tokyo pour lesquelles l'engouement pour les techniques complexes interimages avec compensation de mouvement était moins fort.

L'une des approches prometteuses, largement abordée actuellement, est la définition de codeurs travaillant dans un domaine transformé. Nous étudions actuellement cet axe de recherche au travers de deux approches principales:

- l'utilisation des corrélations intraimages par l'emploi de quantificateur vectoriel sur les blocs transformés,

- l'utilisation des corrélations interimages par la conception de procédure de rafraichissement temporel

rapide des dictionnaires utilisés par la quantification vectorielle.

Cette approche a fournit d'ores et déjà des premiers résultats instructifs notamment concernant les débits excomptés (entre 1 bit et 1,5b/pixel) et les possibilités d'analyse instantanée que fournissent les dictionnaires rafraichis.

REFERENCES BIBLIOGRAPHIQUES

- |1| C. Richard: Estimation réursive de l'état local des contours d'image et application à la prédiction adaptative en codage différentiel des signaux de télévision. **Thèse 3ème cycle, Rennes I, IRISA Rapport n°124, nov.79.**
- |2| F. Kretz: Codage MIC différentiel à prédiction adaptative en télévision:: qualité visuelle, réduction de débit et susceptibilité aux erreurs de transmission. **Annales des Telecom. n°7-8, Juillet-Aout 1982.**
- |3| F. Kretz: Edges in visual scenes and sequences: application to filtering, sampling and adaptive DPCM coding. In "Image Sequence Processing and Dynamic Scene Analysis". Edit. T.S. Huang, 1983.
- |4| A. Benveniste: A measure of the tracking capability of recursive stochastic algorithms with constant gains. **IEEE Trans. on Cont., vol.27, n°3, Juin 1982.**
- |5| A.N. Netravali, J.O. Robbins: Motion-Compensated Television Coding (Part.I). **BSTJ, vol.58, n°3, Mars 1979.**
- |6| J.D. Robbins, A.N. Netravali: Recursive motion compensation: a review. In "Image Sequence Processing and Dynamic Scene Analysis". Edit. T.S. Huang, 1983, NATO ASI Series

- |7| C. Labit: Estimation de mouvement dans une séquence d'images de télévision. **Thèse Docteur-Ingénieur**, Rennes I, Fev. 1982.
- |8| C. Labit, A. Benveniste: Motion estimation in a sequence of television pictures. In "Image Sequence Processing and Dynamic Scene Analysis". Edit. T.S. Huang, 1983.
- |9| P. Robert, C. Labit: Comparative Study between intra-and inter-frame prediction schemes. In: "Image Sequence Processing and Dynamic Scene Analysis". Edit. T.S. Huang, 1983.
- |10| P. Robert: Définition d'un schéma de codage multimodes avec compensation de mouvement pour les séquences d'images de télévision. **Thèse de Docteur-Ingénieur**, Université de Rennes I, nov.83.
- |11| F. Jorgensen: Etude et réalisation d'un processeur temps-réel de compression d'images de télévision. **Thèse de Docteur 3ème cycle**, Univ. de Rennes I.
- |12| A. Habibi: Survey on adaptive image coding techniques. **IEEE Trans. on Com.**, vol.25, n°11, nov.77.
- |13| J.R. Jain, A.K. Jain: Displacement measurement and its application in interframe image coding. **IEEE Trans. on Com.**, vol.29, n°12, Dec. 1981.
- |14| F.J. Santillana Rivero: Etude du codage MICDA de séquences d'images de télévision avec compensation de mouvement. **Thèse de Docteur-Ingénieur**, Univ. de Rennes I, Dec. 85.

A N N E X E S

CONTRIBUTIONS A CES TRAVAUX

1984-1986

A N N E X E I

CONTRIBUTION TO PREDICTIVE CODING AND MOTION ESTIMATION

IPSWICH COST 211 BIS WORKSHOP 7-9 FÉVRIER 1984

CLAUDE LABIT

(PAS D'ACTES)

A N N E X E I I

OPTIMIZATION OF ADDRESSING BIT RATE IN A MULTIMODE CODER WITH MOTION COMPENSATION FOR BROADCAST TELEVISION

J. DEWITTE, C. LABIT, P. ROBERT

(1984 INTERN. ZURICH SEMINAR ON DIGITAL COMMUNICATIONS)

OPTIMIZATION OF ADDRESSING BIT RATE IN A

MULTIMODE CODER WITH MOTION COMPENSATION FOR BROADCAST TELEVISION.

J. DEWITTE (1), C.LABIT (2), P. ROBERT (2)

(1) CNET/CCETT Departement EAS
Rue du Clos Courtel B.P.59
35510 CESSON SEVIGNE FRANCE

(2) IRISA/INRIA Centre de RENNES
35042 RENNES Cédex FRANCE

ABSTRACT :

This study presents a multimode coder scheme with motion compensation. We mainly focalise this paper on algorithms which optimize the addressing bit rate. Using filtering and growing techniques, we hopefully reduce the total number of clusters. Experiments were performed using either videoconferencing sequence or critical television sequence.

Keywords : multimode coder, intra/interframe prediction, motion compensation, segmentation, clustering, filtering, I-D and 2-D addressing.

1. INTRODUCTION.

Using high sampling frequencies for the luminance components of broadcast television sequences (as in CCIR recommendation 601), it appears quite difficult to achieve a high qualite reconstructed signal at low bit rate, only using simple coding algorithms such as non adaptive DPCM or interframe conditional replenishment. The basic structure of the coder considered in this study is a multimode prediction algorithm combining adaptive intrafield prediction, conditional replenishment and interfield motion-compensated prediction.

Such a coding structure provides us both prediction-errors and addressing informations to transmit. Many papers present techniques in order to minimize the prediction-error bit rate. However, the addressing information make use of a large fraction of the total bit rate in conditional coding algorithms.

In this paper, we present methods to encode the addressing informations and describe some simulations to evaluate their performances.

The first step is to decide for each pixel which prediction strategy will be used. This process results in a classification and segmentation of the picture into clusters which falls under one of three headings :

- H ① : fixed areas
- H ② : motion compensated areas
- H ③ : others

Headings ① and ② need only the transmission of their location, ③ needs also the transmission of the pixel's values, coded using an intrafield adaptive prediction DPCM technique. To minimize the number of areas - and thus the addressing information - we present some clustering and growing techniques. Then we attempt to reduce cluster-addressing bit rate either with a 1-line-mnodimensional differential coding of addresses or with a predictive bidimensional coding.

All these enhancements in addressing coding combined with optimized predictors provide us with a significant improvment in terms of total bit rate namely less than 2 bit/pixel for critical natural test sequence of broadcast television type.

The experiments were performed on two test sequences : the first one is the videoconferencing sequence "Judith" (see Fig.1), gracefully given by A. Netravali (bell.Labs), 1 Mhz bandwidth sampled at Nyquist rate, (256 pxels/line) and with little amount of motion (1 or 2 pels/frame) ; the second sequence (see fig.2) is the sequence "Castle" gracefully given by L. Stenger (FTZ, Germany) which involves camera motion (zoom, panning) and object motion (rotation) ; it was sampled at 12 Mhz (624 pixels/line) and often contains quite important motion (3 to 5 pels/frame).

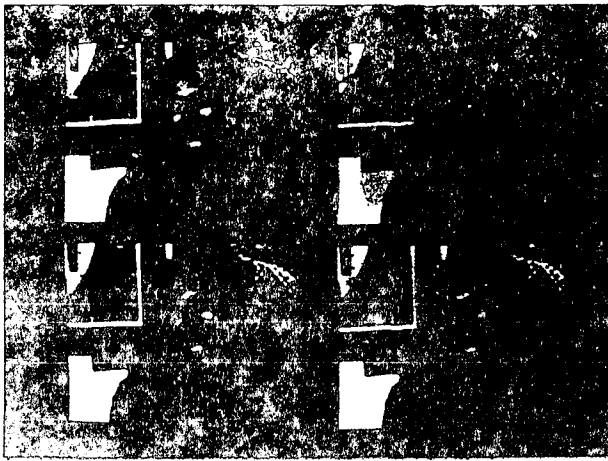


Figure 1

Four consecutive pictures from the videoconferencing sequence "Judith".

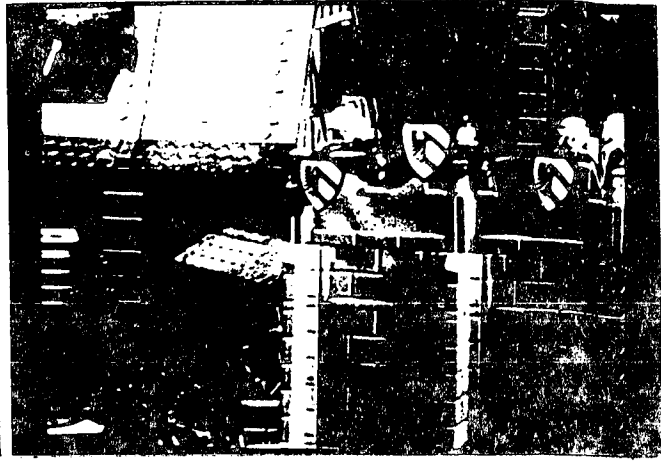


Figure 2

One picture of the T.V sequence "Castle"

II. DESCRIPTION OF THE MULTIMODE CODER :

The study on a multimode coder scheme that we present in this paper, is a follow-up of several researches of our laboratory in the field of intrafield ADPCM coding schemes ([1], [2], [3]) and motion estimation algorithms ([4], [5], [6]). This paper will not describe these algorithms. All details can be found in the references. Briefly we can summarize the approach that we use [7] within all these papers : our approach specifically stresses edges that we consider as stationarity breaks (see Fig.3). Thus we describe algorithms estimating local edge orientation and local edge motion features recursively along spatial edges. All these local observations are filtered along edges to make the algorithms robust to source and channel noises : this process can be described as a state automaton where the inputs are local estimates and the outputs are filtered state components (see Fig.4); the running clock will be line sampling frequency.

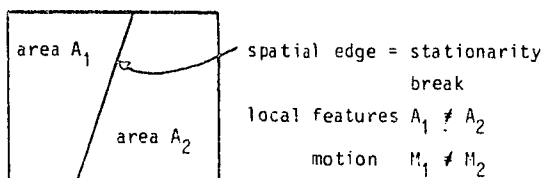
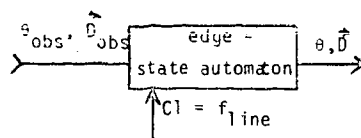


Figure 3

Figure 4
A spatial edge

From all these estimates, an intrafield adaptive prediction and a motion estimation which can be used in motion compensated interframe prediction are derived. We add a simple interframe prediction to these two coding schemes. Thus we obtain all the basic prediction strategies for our multimode coder. Fig.5 shows the general block-diagram of such a coder : three kinds of pixels are involved :

- . fixed pixels : ①
- . motion compensated pixels : ②
- . motion uncompensated pixels : ③

Obviously this classification is not a partition of pixels but we will discuss below how to optimally use the intersections between the three classes.

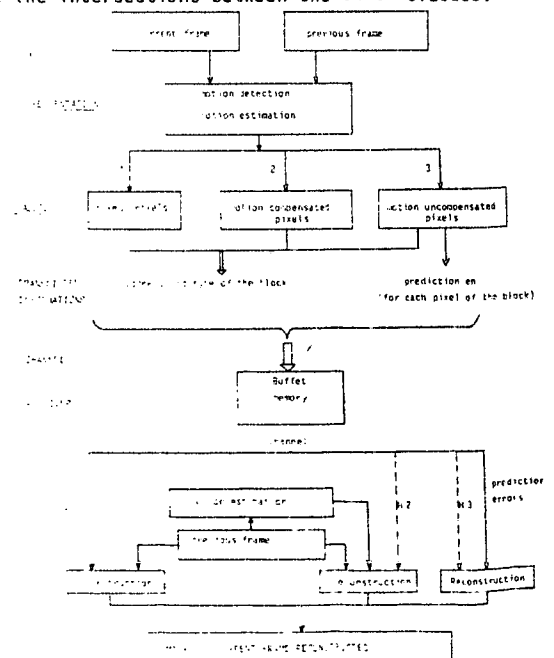


Figure 5 : Block-diagram of the coder

Temporal change detector will be an interframe thresholding and motion estimation will be computed by C. Labit's algorithm [4,5]. All the pixels having the same prediction scheme will be cluster within blocks. Informations about block - addresses and, for the motion uncompensated pixels only, prediction errors have to be transmitted on the channel.

Many papers in the litterature tried to minimize the prediction error bit rate. But it appears that the addressing rate, is quite important into the total bit rate (sometimes nearly 60% of the total bit rate). For this reason, our study was concerned with the re-search of methods to reduce the addressing binary amount. This specific bit rate could be reduced by two means :

1) the minimization of the total number of clusters using some clustering and growing techniques.

2) a well fitted address-coding using a monodimensional or bidimensional differential coding.

We develop both aspects, respectively in section III and IV of this paper.

III . PICTURE SEGMENTATION :

III.1 - Introduction

The first step for such a multimode coder is to decide for each pixel which prediction strategy will be used. This process results in a pixel classification and a segmentation of the picture into clusters. Each pixel will fall under one of three headings :

- heading ① : fixed areas
- heading ② : motion compensated areas
- heading ③ : others.

We can use two temporal flags to perform the pixel classification. These flags are computed using the frame difference (FD) and the displaced frame difference (DFD) :

$$FD = I(x, t) - I(x, t-1)$$

$$DFD(x, \hat{D}) = I(x, t) - I(x - \hat{D}, t-1)$$

We mean by \hat{D} , the predicted motion of the current pixel x between current frame (time t) and previous frame ($t-1$). Several methods have been tested in [8] to predict \hat{D} . We only discuss here a motion compensation scheme using as motion prediction of the current pixel the motion estimation at upper pixel (same pixel in the previous line).

Based on these two temporal values (FD and DFD), we present two segmentation approaches ;

- the first one is a segmentation without filtering and it is based on hierarchical tests on FDs and DFDs.
- the second method attempts to reduce the number of clusters by using an additional heading ④ which contains the intersection subset of pixels ① and pixels ②. This ambiguity (-one pixel could be

declared fixed with a small FD also declared motion - compensated with a small DFD) allows spatial filtering process along the line, and thus clusters pixels in a better efficient way.

III.2 - Hierarchical segmentation :

This segmentation process is shown at Fig. 6. The critical problem consists in the choice of thresholds λ_1 and λ_2 ; they must be defined looking both at objective criteria (noise reduction, segmentation of prediction errors histograms...) and at subjective effects (visual degradation, high quality for reconstructed images...)

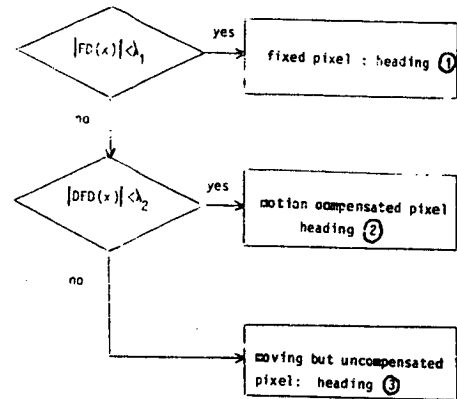


Figure 6 : Hierarchical segmentation

Briefly, λ_1 represents the segmentation threshold between fixed areas and others. It must distinguish true temporal changes from temporal noise (source or quantization noises) ;

λ_2 represents the maximal error that we allow authorize within the reconstruction process of motion-compensated pixels. Above this threshold, the concerned pixel will be refresh using a spatial predictor (algorithm F.Kretz or C.Richard).

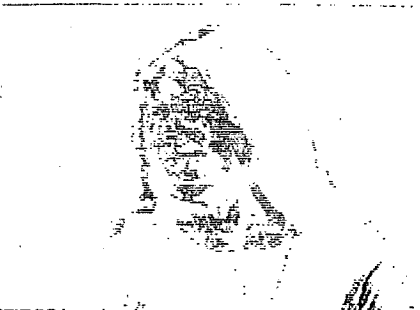
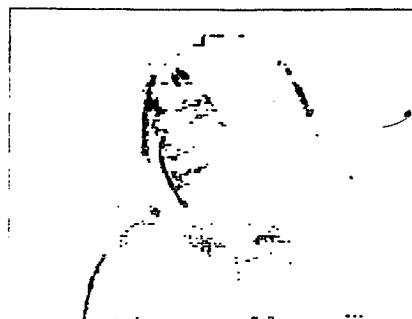
Results : Fig.7 (res. Fig.8) shows the segmentation of image "Judith" (resp. image "Castle"). into three sub-images : pixels with heading ①, ② and ③. We also compare edge pictures and show, by this way, the close correlation between edges and borders of the different clusters.

First conclusion : pixels with heading ③ are mainly located on moving edges, i.e on spatial and temporal stationarity breaks, but also on moving textured areas i.e high local-gradient activity areas. The involved segmentation process resulting in these subimages creates too many clusters (especially in sequence "Castle") with little sizes. To improve the clustering process, some new techniques must be implement : we suggest in the following section to define a true partition of the image between disjonctive subsets and a subsequent filtering (along a line).

Pixels under heading ①



Pixels under heading ③



Pixels under heading ②

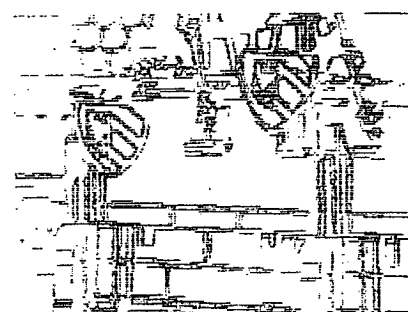
Edge pixels

FIGURE 7 : Segmentation of image "Judith" using hierarchical segmentation algorithm . Comparison with the edge picture.

Pixels under heading ①



Pixels under heading ③



Pixels under heading ②

Edge pixels

FIGURE 8 : Segmentation of image "Castle" using hierarchical segmentation algorithm. Comparison with the edge picture.

III.3 - Segmentation with spatial filtering :

III.3 a - Segmentation

We will consider in this part a segmentation of the image into four disjunctive classes :

- .heading ① : fixed and motion-uncompensated pixels
- .heading ② : displaced and motion compensated pixels
- .heading ④ : fixed and motion compensated pixels
- .heading ③ : others.

Class ④ is an additionnal but temporary subset because, all the pixels ④ will fall either under heading ① or ② as a result of the filtering process

Heading ④ $|FDI| < \lambda_1$ and $|DFDI| < \lambda_1$
 Heading ① $|FDI| < \lambda_1$ and $|DFDI| \geq \lambda_1$
 Heading ② $(|FDI| \geq \lambda_1 \text{ and } |DFDI| < \lambda_1)$
 or $(\lambda_1 \leq |DFDI| < \lambda_2 \text{ and } |FDI| \geq |DFDI|)$
 Heading ③ others....

Figure 9 : Partition within 4 disjonctive classes.

Fig. 9 presents the segmentation algorithm and shows that, unlike hierarchical segmentation of III.2 , it sets and tests the temporal flags FD and DFD in a parallel way.

Results : Fig. 10 and 11. show the partition of pixels between the four subimages.

III.3.b - Filtering process of clusters :

Our analysis is to consider that we can use the ambiguity of pixels with heading ④ to select the best prediction strategy between either heading ① or ② to minimize the total number of cluster transitions without introducing any impairment . The filtering process consists in a spatial hysteresis scheme ; several methods in [8] have been compared ; we only present the optimal one.

The filtering process for pixels which fall under heading ④ can be expressed as :

- If X_{n-1} ⑥ heading H ① X_n ⑥ heading H ①
- if X_{n-1} ⑥ heading H ② X_n ⑥ heading H ②
- if X_{n-1} ⑥ heading H ③ X_n ⑥ heading H ②

Intentionally, we magnify, the number of pixels with heading ②, to reduce the "dirty window" impairment issued from a rough temporal thresholding.

Results : Fig. 12 (resp.13) gives results when this filtering is done on sub-images of fig. 10 (resp Fig.11) If clearly appears that the number of clusters is significantly reduced (see III.5).

III.4 - Some improvements

Even if a filtering is involved, too many little clusters (clusters with only few pixels) remain in the segmented picture. We often can expect that it would be more fruitful to transmit few additionnal prediction

errors (by setting the heading ③ to some clusters ① or ②) than to transmit both informations of the prediction heading and cluster address which are quite expensive in terms of bit rate for the smallest clusters.

This "sweeping step" could be performed in two ways :

* an "objective" sweeping rule which select the optimal heading of prediction for the current cluster by comparison of bit rates computed either without or with sweeping process ; it involves the propagation of a 4×1 mask and creates a maximum delay of two pixels for the knowledge of the prediction heading after sweeping process. This processing does not create any impairment within the reconstructed picture.

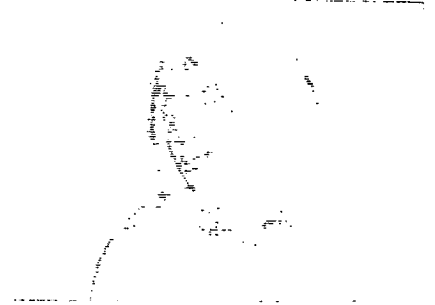
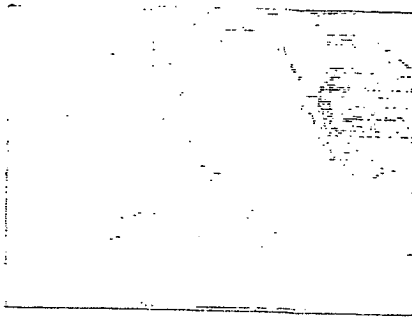
* a "subjective" sweeping rule can be performed. In this case, the constraint is to sweep and roughly erase all the isolate clusters and replace them according to the adjacent cluster headings. This sweeping process hopefully does not create any visible impairment and must be declared available after visual tests.

III.5 - Results (fig. 14)

All the experiments are described in [8]. We only present some comparative results with two sets of thresholds (λ_1 and λ_2) for both sequence "Judith" and "Castle". Results are expressed in terms of number of clusters per line and the percentage of pixels in each prediction class ①, ②, or ③. We can observe a reduction factor nearly about 3.5 from the initial number of clusters.

	THRESHOLDS		ALGORITHM OF SEGMENTATION	NUMBER OF CLUSTERS/line	PERCENTAGE OF PIXELS		
	λ_1	λ_2			①	②	③
"JUDITH"	4	7	Hierarchical segmentation	19.5	88	9.5	2.5
			Segmentation with filtering	8.7	51.2	46.3	2.5
			Segmentation with filtering and sweeping	5.5	51	46	3
	3	5	Hierarchical segmentation	27	84	12	4
			Segmentation with filtering	14	50.5	45.5	4
			Segmentation with filtering and sweeping	8	50	45	5
"CASTLE"	7	12	Hierarchical segmentation	190	47	36	17
			Segmentation with filtering	103.6	13	70	17
			Segmentation with filtering and sweeping	52.5	11	68	21
	5	9	Hierarchical segmentation	229	37	40	23
			Segmentation with filtering	144	13	64	23
			Segmentation with filtering and sweeping	68	10	61	29

FIGURE 14



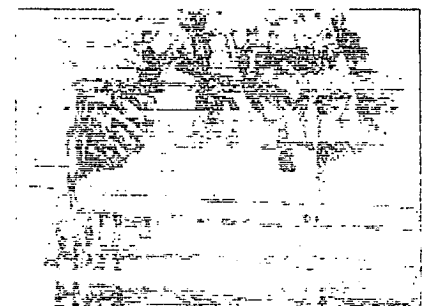
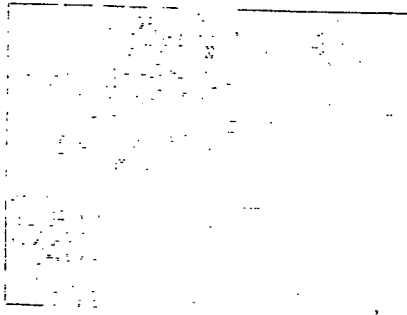
Pixels under heading ④

Pixels under heading ③

FIGURE 10 : Segmentation of image "Judith" within four disjonctive classes.

Pixels under heading ①

Pixels under heading ②



Pixels under heading ④

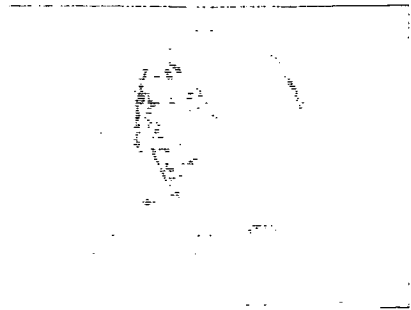
Pixels under heading ③

FIGURE 11 : Segmentation of image "Castle" within four disjonctive classes.

Pixels under heading ①



Pixels under heading ③



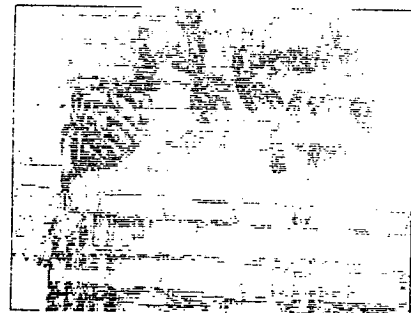
Pixels under heading ②

FIGURE 12 : Segmentation with filtering of image "Judith" and pixel clustering under three headings.

Pixels under heading ①



Pixels under heading ③



Pixels under heading ②

FIGURE 13 : Segmentation with filtering of image "Castle" and pixel clustering under three headings.

IV. ADDRESS CODING

The picture segmentation provides us with three subimages (headings ①, ②, ③) which agree with each possible strategy of prediction. These subimages are defined on each picture line with clusters. We have to transmit several informations for each cluster :

* The corresponding heading ①, ②, or ③: two bits are needed but one of the four code words are not used.

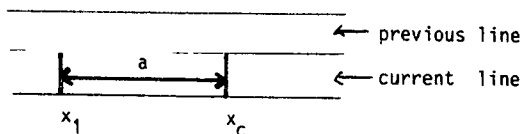
* The address of the cluster : we will use some address-coding methods already used about digital facsimile coding application.

IV.1 - Addressing algorithms

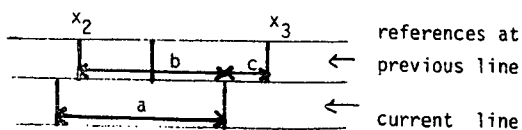
All the following algorithms involve differential coding techniques and will be well protected against transmission errors. Two different addressing methods are intended (see Fig.15)

. a 1 - D differential addressing only using the border of the last cluster at the current line

. a 2 - D differential addressing using either the last cluster address at the current line or any cluster reference from the previous line. This cluster reference could be associated with a cluster border or a spatial edge. In this case, three differential addresses could be concurrently transmitted : relative addresses a, b, c. When address c is selected, an additional information about the number of jumps over clusters at previous line must be together transmitted. Obviously, an optimal compromise has to be found between all the different transmitted informations (addresses, headings, jumps,...)



1-D addressing coding: the transmitted relative address is: $a = \text{add}(x_c) - \text{add}(x_1)$



2-D addressing coding:

x_1 : last cluster address: $a = \text{add}(x_c) - \text{add}(x_1)$

x_2 : first address on the right of x_1 on previous line $b = \text{add}(x_c) - \text{add}(x_2)$

x_3 : nearest reference from x_c at previous line $c = \text{add}(x_c) - \text{add}(x_3)$

FIGURE 15

IV.2 Brief results and conclusion

Briefly, we present the addressing algorithms which were selected after reduction rate valuation (in terms of address entropy). We have to distinguish the different available transitions: ① → ②, ② → ①, ② → ③ or ③ → ② or ③ → ①

the addresses for transitions ③ → ② are not transmitted or ①

ted by using a specific codeword for ends of blocks of prediction errors.

Transitions ① → ② or ② → ① have bad correlation from line to line and a 1-D relative addressing appears quite competitive. However for transition ③ → ③, 2-D addressing is more efficient. Using binary entropy criterion, we obtain improvements nearly about 0.9b/add and 1.7b/add for respectively images "Castle" and "Judith" between 1-D addressing and 2-D addressing (transmitted address = c) algorithms.

In terms of conclusion, we can expect that all these addressing coding enhancements combined with optimized prediction strategies (especially with motion compensation scheme) allow to reduce quite significantly the total bit rate namely less than 2bits/pixel.

REFERENCES

- [1]: C. Richard, Estimation recursive de l'état local des contours d'image et appl. à la prédiction adaptative ADPCM en télévision. IRISA Report n°24 (Nov 79)
- [2]: C. Richard, A. Benveniste, F. Kretz : Recursive estimation of local characteristics of edges in T.V pictures; as applied to ADPCM coding, Report CCETT (to be published in IEEE on COMM.)
- [3]: J. Dewitte, Détection et suivi de contours en codage MICD Thesis INSA Rennes Dec 81
- [4]: C. Labit, A. Benveniste, Motion estimation in T.V sequences, NATO ASI Series (editor HUANG): Image sequence Processing and dynamic scene analysis (83)
- [5]: C. Labit, Estimation de mouvement dans une séquence d'images de télévision. Thesis IRISA Rennes (Feb 82)
- [6]: A. Netravali, J.O Robbins, Motion-compensated Television coding BSTJ Vol 58 n°3 (March 1979)
- [7]: F. Kretz, Edges in Visual scenes and sequences Application to filtering, sampling and ADPCM coding same ref. as [4].
- [8]: P. Robert, Définition d'un schéma de codage multimodes avec compensation de mouvement pour les séquences d'images de Télévision Thesis IRISA (Nov 83)

A N N E X E I I I

CODAGE ADAPTATIF MULTIMODES POUR LES SEQUENCES D'IMAGES DE TELEVISION

P. ROBERT, C. LABIT

Traitement, Synthèse, Technologie et Applications

BIARRITZ - Mai 1984 -

CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION

Adaptive multimode coding scheme for broadcast television sequences .

PHILIPPE ROBERT CLAUDE LABIT

IRISA/CENTRE INRIA DE RENNES · CAMPUS DE BEAULIEU, AV. DU GÉN. LECLERC 35042 RENNES CEDEX FRANCE

RESUME

Afin de réduire le débit binaire d'informations à transmettre pour le codage des séquences d'images de Télévision Haute Qualité ($f_s=13.5\text{Mhz}$), cette étude présente un schéma de codage MICDA, à débit variable, et à stratégies de prédiction multiples.

Les prédicteurs sont choisis intrimages, ou inter-images avec ou sans compensation de mouvement. Ces trois types de prédicteurs conduisent :

- . à une optimisation locale en chaque pixel de l'erreur de prédiction qui n'est transmise que pour les points non compensables en mouvement;
- . à l'envoi d'informations d'adresses des différentes zones de prédiction qui représentent bien souvent une part importante du débit binaire total de transmission.

Des techniques optimisées d'adressage ont déjà été proposées dans le domaine du codage par plages, du codage d'images Télécopie...etc. Celles-ci méritent d'être étendues dans le cadre général d'un codeur multimodes.

Pour ce faire, cette étude propose deux axes de recherche pour la réduction du débit d'adresses. La première optimisation porte sur le nombre total d'adresses qui peut être réduit par des algorithmes de segmentation d'images et de regroupements de paquets. La deuxième a été menée grâce à une étude comparative sur les algorithmes d'adressage différentiel monodimensionnel ou bidimensionnel. La notion de prédiction d'adresses permet de réduire efficacement la taille des adresses différentielles à coder.

L'ensemble de ces techniques de prédiction et d'adressage a été testé sur deux séquences d'images (à caractéristiques différentes) et autorise pour chacune d'elles des gains significatifs en compression de l'information à transmettre.

MOTS CLES : Codage, séquences, prédiction intra/inter, compensation de mouvement, adressage différentiel.

SUMMARY

This paper presents an ADPCM multimode coder scheme reducing the transmission bit rate of high quality ($f_s=13.5\text{Mhz}$) television picture sequences.

Set of predictors have been used in both intra- and interframe domains. Motion compensation has been performed to improve the coding of moving areas. This multimode structure authorizes :

- . a local optimization, at each pixel, of the prediction error which is only transmitted for motion-uncompensated pixels.
- . an addressing bit rate which is often quite important into the total bit rate and has to be locally optimized.

Some techniques have been already used for block coding, telecopy... and could be extended to any multimode coding scheme.

We present two studies reducing the addressing bit rate. The first step is to reduce the total number of addresses. It can be performed using segmentation, filtering and growing algorithms applied to the image clusters. Then some differential (1-dimensional or 2-dimensional) addressing coding can be helpful to optimally code these boundaries.

All of these coding algorithms have been tested using sequences with different spatial and temporal characteristics and enable quite significant decrease in terms of information amount to be transmitted.

KEYWORDS : Image sequence coding, intraframe, interframe prediction, motion compensation, differential addressing

CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION
Adaptive multimode coding scheme for broadcast television sequences .
PHILIPPE ROBERT CLAUDE LABIT

I . INTRODUCTION

Les études que nous avons menées ([1], [2], [3]) dans le domaine du codage de séquences d'images de télévision visent de manière prospective à cerner, analyser et utiliser les redondances quelles que soient leurs natures (spatiales, temporelles, spatio-temporelles) présentes dans une séquence d'images. Pour ce faire, il convient de considérer une séquence d'images comme un signal spatiotemporel $I(x,y,t)$ où les corrélations sont à rechercher simultanément dans les trois dimensions:

- i) l'utilisation des corrélations spatiales (en x et y) autorise des prédictions intraimages bidimensionnelles efficaces; cette possibilité doit être utilisée dans les zones spatialement homogènes;
- ii) l'analyse des corrélations temporelles permet de segmenter l'image en zones fixes et zones en mouvement. Deux types d'algorithmique à complexité différente peuvent être introduits à ce niveau: une simple détection de mouvement permet de supprimer la transmission des informations d'image contenues dans les zones fixes; pour les zones en mouvement, une étape supplémentaire—celle d'une estimation quantitative des mouvements des points "changés" temporellement—convient d'être introduite afin de prédire par compensation de mouvement les points d'images des zones en mouvement.

Les algorithmes ont été testés sur deux séquences d'images à caractéristiques spatiales et temporelles différentes (Fig 1 et 2)

II APPROCHE ET DEFINITION DE LA STRUCTURE DU CODEUR

II.1 : Approche adoptée

La nature prospective de l'étude présentée ici nous a imposé la volonté d'extraire et d'utiliser au maximum

les redondances présentes dans l'image et nous a conduit à opter pour un système de codage à débit variable où aucune information redondante du signal d'image n'est transmise. Le caractère variable du débit est créé par :

- i) un envoi non systématique d'informations d'erreurs; une erreur de prédiction n'est transmise que pour les points non compensables en mouvement;
- ii) un codage statistique (code de Huffman par ex.) des informations transmises.

Dans ce cadre algorithmique, deux volumes principaux d'informations à transmettre sont à minimiser afin de réduire le débit total de transmission :

- 1) la minimisation du débit d'erreurs consiste en la recherche de prédicteurs optimisés. De nombreuses recherches portent sur ce sujet ([4], [5], [6]). Nous présentons ici un schéma de prédiction adaptatif multimodes à prédicteurs tantôt intraimages, tantôt interimages avec ou sans compensation de mouvement.

- 2) un codage non systématique induit l'introduction d'adresses de zones d'image à stratégies de prédiction distinctes. Ce volume d'informations d'adressage représente bien souvent une part importante du débit total binaire (jusqu'à 60% de ce débit [7]). Peu d'études en codage multimodes ont cherché à minimiser ce débit d'adresses par des techniques prédictives de codage d'adresses. De telles techniques optimisées d'adressage ont été proposées dans le domaine du codage par plages, du codage d'images Télécopie...Celles-ci méritent d'être étendues aux structures multimodes. De telles techniques d'adressage suppose une protection renforcée des codes-canal choisis vis-à-vis des erreurs de transmission.

Figure 1 : 4 images consécutives de la séquence "JUDITH" de visiophone ($f_{\text{éch}}=2\text{Mhz}$, mouvements jusqu'à 2 pixels/fr)

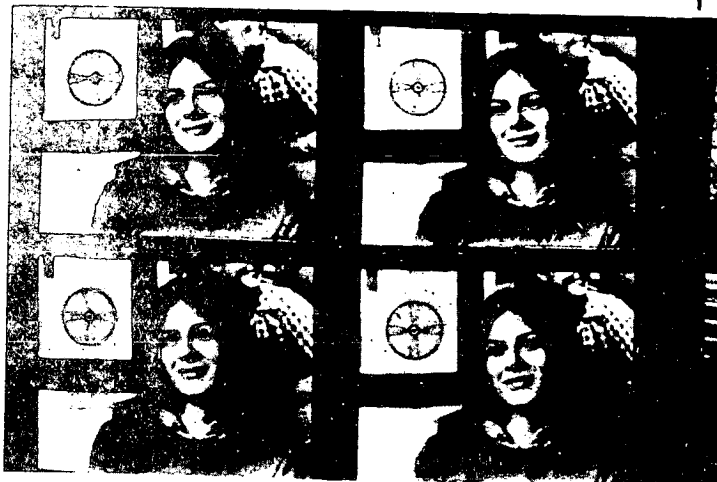
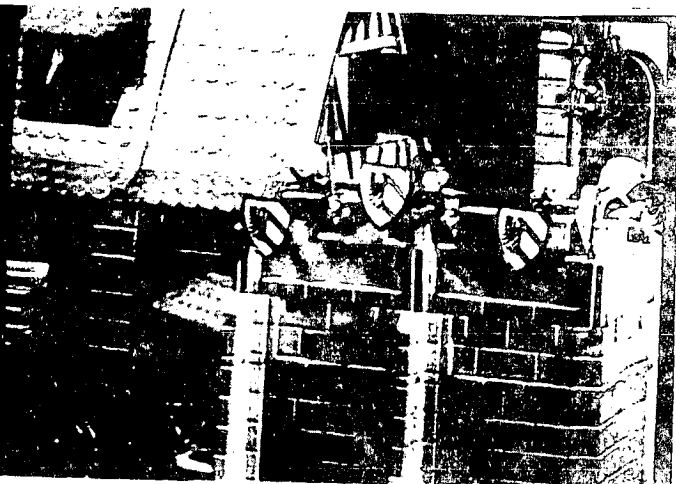


Figure 2 : Image source de la séquence "CHATEAU" ; ($f_{\text{éch}}=12\text{Mhz}$, mouvements jusqu'à 6 pixels/frame)



Nous ne détaillons pas ici les algorithmes choisis pour la définition des différents modules de prédiction mais précisons quelles ont été les stratégies de prédiction finalement retenues pour une minimisation du débit de transmission d'erreurs.

Ce module consiste à rechercher dans l'image précédente le point d'image situé à la même localisation spatiale que le point courant à coder. Dans ce cas, si P désigne le prédicteur et x_c le pixel à coder :

II.2b : module de prédiction interimage avec compensation de mouvement M_2

Toutefois ces algorithmes convergent lentement (ou parfois divergent) lors de franchissements de non-stationnarités spatiotemporelles (contours en mouvement). Pour pallier à ce manque, il est nécessaire de construire une procédure d'estimation de mouvement spécifique pour de telles zones ruptures spatiales ([1], [2]). Ces zones sont situées sur les points contours.

$P(x_c, t) = \hat{I}(x_c - \hat{D}_{x_c}, t-1)$, \hat{D}_{x_c} désignant la prédiction de mouvement au point x_c .

* Module M_3 : si x est prédit comme point non contour } $P(x_c) = \frac{A+C}{2}$
 • ou contour $\theta \geq 0$
 • contour $\theta < 0$ } $P(x_c) = \frac{A+D}{2}$

B	C	D
A	x_C	

↖ point à prédire

Comme pour l'estimation de mouvement, l'idée sous-jacente vise à définir des prédicteurs adaptés au contexte local de l'image. Des algorithmes de détection (seuillage sur gradients spatiaux) et de suivi (modèle et représentation d'état des contours) de contours ont été définis. Ils permettent d'estimer divers paramètres d'états de contours (longueur, épaisseur, orientation...) et d'adapter ainsi les prédicteurs en fonction notamment de l'orientation des contours. Différents lots de prédicteurs adaptatifs ont été testés [6]. L'étude décrite ici a utilisé un schéma adaptatif à 3 prédicteurs commutés, comme le résume la Fig. 3.

La Figure 4 montre le schéma général du codeur étudié. Ce codeur vise à segmenter l'image en trois types de points :

- 1) les points "fixes" pour lesquels le module de prédiction interimage simple M_1 sera utilisé;
- 2) les points "compensés en mouvements" qui seront prédits par compensation de mouvement (module M_2);
- 3) les points restants étant prédits par le module M_3 dans le domaine intrimage. Pour eux seuls, seront transmises les erreurs de prédiction quantifiées.

Une telle structure multimodes de codage induit donc la définition d'algorithmes de segmentation et classification de l'image en trois types de zones à stratégies de prédiction différentes. Cette partition de l'image est à optimiser du point de vue nombre de zones et codage d'adresses de ces zones.

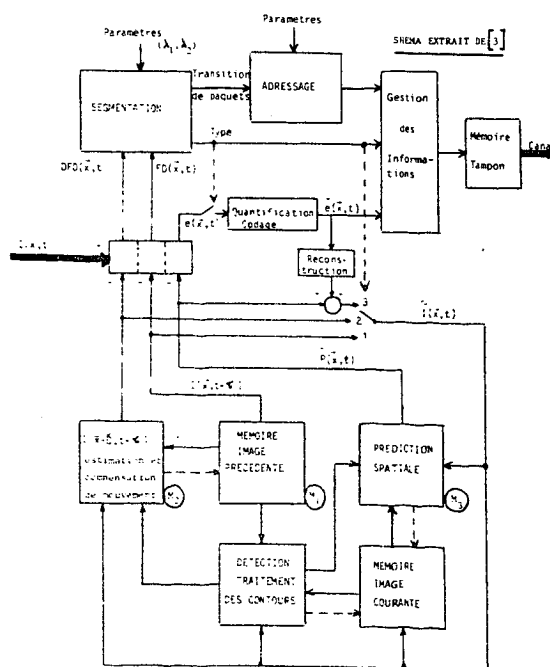


FIGURE 4 : Schéma général du codeur

CODAGE ADAPTATIF MULTIMODES POUR LES SEQUENCES D'IMAGES DE TELEVISION
Adaptive multimode coding scheme for broadcast television sequences .
PHILIPPE ROBERT CLAUDE LABIT

III . SEGMENTATION ET CLASSIFICATION DE L'IMAGE

III . 1 : Algorithme de segmentation hiérarchique.

* La première méthode de segmentation proposée est effectuée à partir des deux données temporelles suivantes disponibles en chaque pixel :

- la différence interimage ("Frame difference = FD")

$$FD(x,t) = I(x,t) - \hat{I}(x,t-1)$$

- la différence interimage avec compensation de mouvement ("Displaced Frame Difference = "DFD")

$$DFD(x,t) = I(x,t) - \hat{I}(x-\hat{D}(x,t), t-1)$$

* L'algorithme de segmentation est décrite en Figure 5. Il consiste à détecter tout d'abord les 3 points fixes, puis les points compensés et enfin les points restants, critiques pour la prédiction, qui sont principalement situés sur les zones contours. Les résultats de cette segmentation sont visualisés en Annex Figure A . Le choix des seuils de segmentation λ_1 et λ_2 doit être effectué en respect de critères objectifs (réduction du bruit temporel, segmentation non redondante) mais également de critères subjectifs (dégradation visuelle maximale autorisée, haute qualité après reconstruction...) Le seuil λ_1 discrimine les zones fixes des zones en mouvements mais aussi les "vrais" changements temporels du bruit temporel de source ou de quantification. Une valeur élevée de λ_1 permet une élimination de celui-ci mais induit des effets dynamiques "de vitre sale" fortement nuisibles. Le seuil λ_2 est à considérer comme la valeur maximale de l'erreur temporelle autorisée dans les zones à mouvement compensé.

III . 2 : Segmentation avec filtrage spatial.

III.2.a : Certains points d'images, au vu des deux paramètres temporels FD et DFD, peuvent être déclarés à la

fois fixes et compensés en mouvement. Cette remarque préliminaire amène tout naturellement l'introduction d'une classe supplémentaire dans le processus de segmentation de l'image. L'ajout de cette nouvelle classe "temporaire" permettra de tenir compte de l'ambiguïté inhérent à certains pixels et de partitionner l'image en quatre classes disjointes dont les définitions (avec notations de la Figure 5)

$$\text{classe } T'_4 = T_1 \cap T_2$$

$$\text{classe } T'_1 = T_1 - T_2$$

$$\text{classe } T'_2 = T_2 - T_1$$

$$\text{classe } T'_3 = T_3$$

III.2.b : Filtrage spatial sur la classe T'_4 . Plusieurs méthodes de filtrage sur les pixels de type T'_4 ((3)) ont été testées. La plus efficace semble être un schéma d'hystérésis spatial ou politique conservatrice d'affectation de type entre le point précédent et le point courant :

si X_n (point à classer) $\in T'_4$
alors si X_{n-1} (point précédent) $\in T'_1$ alors $X_n \in T'_1$
si $X_{n-1} \in T'_2$ alors $X_n \in T'_2$
si $X_{n-1} \in T'_3$ alors $X_n \in T'_2$

III.2.c : Résultats et améliorations. Les résultats de segmentation sont visualisés en Annexe. Figure B montre la partition en 4 classes disjointes avant filtrage sur une des séquences. Figures C et D montrent la partition finale en 3 classes après filtrage par hystérésis pour les deux types de séquences utilisées.

De nombreuses améliorations en termes de segmentation optimale pour une qualité d'image accrue et une minimisation du nombre d'adresses de paquets à transmettre. D'une part des algorithmes d'hystérésis à mémoire permettent de parfaire le filtrage. D'autre part il est nécessaire d'introduire des règles de nettoyage supprimant les paquets de trop faible taille. Ces nettoyages peuvent être effec-

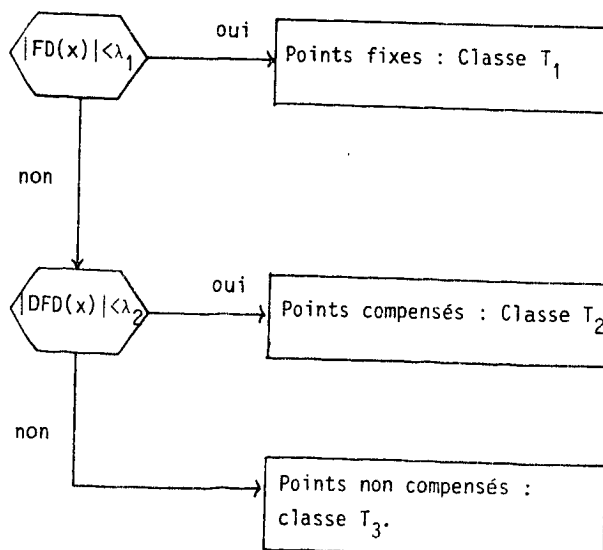


FIGURE 5

Algorithme de segmentation
hiérarchique en 3 classes

CODAGE ADAPTATIF MULTIMODES POUR DES SÉQUENCES D'IMAGES DE TÉLÉVISION
Adaptive multimode coding scheme for broadcast television sequences .
PHILIPPE ROBERT CLAUDE LABIT

tués sans dégradations pour l'image reconstruite (remplacement d'adresses par des erreurs de prédiction à transmettre) ou avec dégradations (pontage de paquets) mais non perceptibles par l'oeil.

III . 3 : Résultats et conclusions.

Nous présentons en Figure 6 les résultats des différents algorithmes sur les deux séquences avec pour chacune d'entre elles, deux jeux de seuils λ_1 et λ_2 . Il apparaît nettement que les techniques de filtrage et nettoyage permettent de fortes réductions du nombre d'adresses à transmettre et une meilleure répartition des points dans les trois classes initiales en privilégiant celle des points compensés en mouvement (minimisation des effets de "vitre sale" dynamiques).

SÉQUENCES	SEUILS	ALGORITHME DE SEGMENTATION	NOMBRE D'ADRESSES PAR LIGNE	POURCENTAGE DES POINTS
"MOTION"	1 2	Segmentation hiérarchique	15,5	68 9,5 2,5
		Segmentation avec filtrage	8,7	51,2 46,3 2,5
		Segmentation avec filtrage et nettoyage	5,5	51 46 3
	1 3	Segmentation hiérarchique	27	54 12 4
"MOTION"	1 3	Segmentation avec filtrage	14	55,5 45,5 4
		Segmentation avec filtrage et nettoyage	8	55 45 5
		Segmentation hiérarchique	100	47 36 17
	1 2	Segmentation avec filtrage	103,6	15 70 17
"MOTION"	1 2	Segmentation avec filtrage et nettoyage	10,5	11 58 21
		Segmentation hiérarchique	129	37 40 23
		Segmentation avec filtrage	114	13 64 23
	1 3	Segmentation avec filtrage et nettoyage	68	17 61 29

FIGURE 6 : Segmentation de l'image

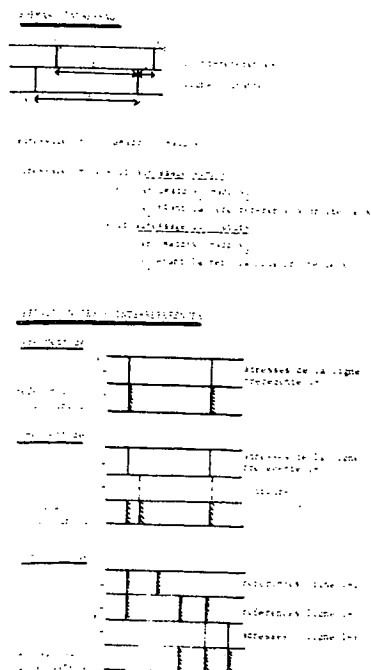


FIGURE 7 :
Adressage
bidimensionnel

IV ALGORITHMES D'ADRESSAGE :

Pour réduire le débit binaire résultant du codage des adresses des zones précédemment segmentées, il convient également de prendre en compte les corrélations existantes entre adresses à coder par un adressage différentiel monodimensionnel (sur une ligne d'image) ou bidimensionnel (sur plusieurs lignes). Certaines techniques d'adressage de ce type ont déjà été développées dans d'autres domaines (télécopie, fac-similé..) et méritent d'être rappelées ([8]) :

IV.1 : Adressage différentiel bidimensionnel

Il s'agit d'adresser les zones différenciellement sur la ligne courante et également en comparaison avec des références de zones de la ligne précédente :

..La méthode EDIC vise à détecter trois types d'événements: i) une apparition d'adresse pour laquelle aucune adresse ne correspond à la ligne précédente

ii) une disparition d'adresse par rapport à une adresse existante à la ligne précédente

iii) un déplacement borné d'adresses entre les deux lignes. Chaque type sera caractérisé et codé différemment.

..La méthode RAC choisit en chaque adresse à coder $add(x_c)$ la stratégie de codage minimale en distance entre l'adressage différentiel 1-D et 2-D. Un indicateur de permutation de mode d'adressage est alors transmis.

..La méthode READ combinant les deux méthodes précédentes apparaît comme un système efficace d'adressage pour le codage des documents.

Nous avons repris et généralisé ces méthodes d'adressage bidimensionnel dans le cadre d'un schéma de codage multimodes.

IV.2 : Application au codeur multimodes :

Avec les notations de la Figure 7, nous distinguerons l'adressage différentiel 1-D qui transmet

$a_1 = add(x_c) - add(x_1)$, x_1 étant la dernière adresse transmise

et l'adressage différentiel 2-D qui transmet

$a_2 = add(x_c) - add(réf)$, $réf$ étant l'adresse-pixel de référence par rapport auquel l'adresse courante $add(x_c)$ va être calculée.

De nombreuses simulations [3] ont été faites pour tester les diverses définitions possibles de points-références. Trois schémas principaux (à complexité croissante) peuvent être retenus :

i) les point-références d'adresses pour la ligne courante sont les adresses de paquets de la ligne précédente

ii) A ces références peuvent être adjointes les adresses des points contours détectés à la ligne précédente. Cette amélioration est introduite pour tenir compte de

CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION Adaptive multimode coding scheme for broadcast television sequences . PHILIPPE ROBERT CLAUDE LABIT

la forte corrélation existante parfois entre les points de transition de zones et les points contours (voir la Fig. A en Annexe). La prise en compte de l'orientation des contours dans cette projection n'apporte qu'un faible gain eu égard à la complexité algorithmique supplémentaire

iii) Une définition récursive de ces points-références par extension de la mémoire des points de références à la ligne ante-pénultième est décrite en Fig. 7. Cet algorithme facilement implantable permet des adressages efficaces même en présence de trous dans les corrélations ligne-à-ligne des adresses de zones.

L'adressage différentiel bidimensionnel pourra être défini selon deux types :

1) l'adressage par $b = \text{add}(x_c) - \text{add}(x_2)$, simple en réalisation, est sous-optimale du point de vue distance référence-adresse courante.

2) l'adressage avec sauts de références permet de minimiser les adresses différentielles mais ajoute des informations annexes (nombre de sauts de références) à transmettre

IV.3 : Résultats et Conclusion :

Les résultats pour les deux séquences (Figure 8) sont exprimés en entropie théorique calculée sur l'histogramme d'adresses et en pourcentage de répartition entre l'adressage mono- et bidimensionnel. La nature des transitions de zones est distinguée également au vu de la forte disparité de corrélation ligne-à-ligne entre les différents types de transitions (voir Annexe).

Il apparaît ainsi que l'introduction d'un codage d'adresses bidimensionnel n'est fortement efficace que dans les zones fortement corrélées verticalement, comme se trouvent être les zones de points non compensés de type T_3 . Dans le cas de séquence à mouvement peu complexe, le gain de l'adressage différentiel 2-D est toutefois sensible sur tous les types de transitions.

CONCLUSION :

L'ensemble des techniques algorithmiques présentées ici ont cherché, tant au niveau prédiction de l'intensité lumineuse qu'au niveau adressage de zones, à utiliser les redondances (corrélations spatiales et temporelles) pour réduire le débit binaire d'informations à transmettre. De telles techniques permettent d'espérer atteindre des débits faibles (avoisinant les 2bit/pixel) pour des séquences de Télévision complexes et de haute qualité.

	TYPE DE TRANSITION	MODE D'ADRESSAGE	ENTROPIE	REPARTITION EN %	
				1-D	2-D
JUDITH	transition $T_1 \rightarrow T_2$ ou $T_2 \rightarrow T_1$	monodimensionnel	5,7	100	
		bidimensionnel simple	5,0		
		bidimensionnel avec sauts	5,0	48	52
	transition $X \rightarrow T_1$	monodimensionnel	5,4	100	
		bidimensionnel simple	3,8		
		bidimensionnel avec sauts	4,5	32	68
CHATEAU	transition $T_1 \rightarrow T_2$ ou $T_2 \rightarrow T_1$	monodimensionnel	6,8	100	
		bidimensionnel simple	6,7		
		bidimensionnel avec sauts	6,7	66	34
	transition $X \rightarrow T_1$	monodimensionnel	5,4	100	
		bidimensionnel simple	4,4		
		bidimensionnel avec sauts	4,4	36	64

FIGURE 8

REFERENCES :

- [1] : C. LABIT : Estimation de mouvement dans une séquence d'images de télévision
Thèse de Docteur-Ingénieur Rapport IRISA n°18/ Fév. 82
- [2] : C. LABIT, A. BENVENISTE : Motion estimation in T.V sequences
NATO ASI Series (édité par T.S HUANG): Image Sequence Processing and dynamic scene analysis (83)
- [3] : P. ROBERT : Contribution à la définition d'un schéma de codeur multimodes avec compensation de mouvement
Thèse Docteur-Ingénieur IRISA Nov. 83
- [4] : A.N NETRAVALI, J.O LIMB : Picture Coding : a review
Proceedings of the IEEE Vol 68 n°3 Mars 80
- [5] : A.K JAIN : Image data compression : a review
Proceedings of the IEEE Vol 69 n°3 Mars 81
- [6] : F. KRETZ : Edges in visual scenes and sequences. Application to filtering, sampling and ADPCM coding.
Même référence que [2]
- [7] : A.N NETRAVALI, J.O ROBBINS : Motion compensated Television Coding
BSTJ Vol 58 n°3 Mars 79
- [8] : Y. YASUDA : Overview of digital Facsimile Coding Techniques
Proceedings of IEEE Vol 68 n°7 July 80

CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION
Adaptive multimode coding scheme for broadcast television sequences.

PHILIPPE ROBERT CLAUDE LABIT

ANNEXE (L'ensemble de ces résultats graphiques sont
extrait de la référence 3)

FIGURE A : Partition de l'Image "JUDITH" en 3 classes suivant l'algorithme de segmentation hiérarchique
Pour comparaison, est visualisée également l'image contours

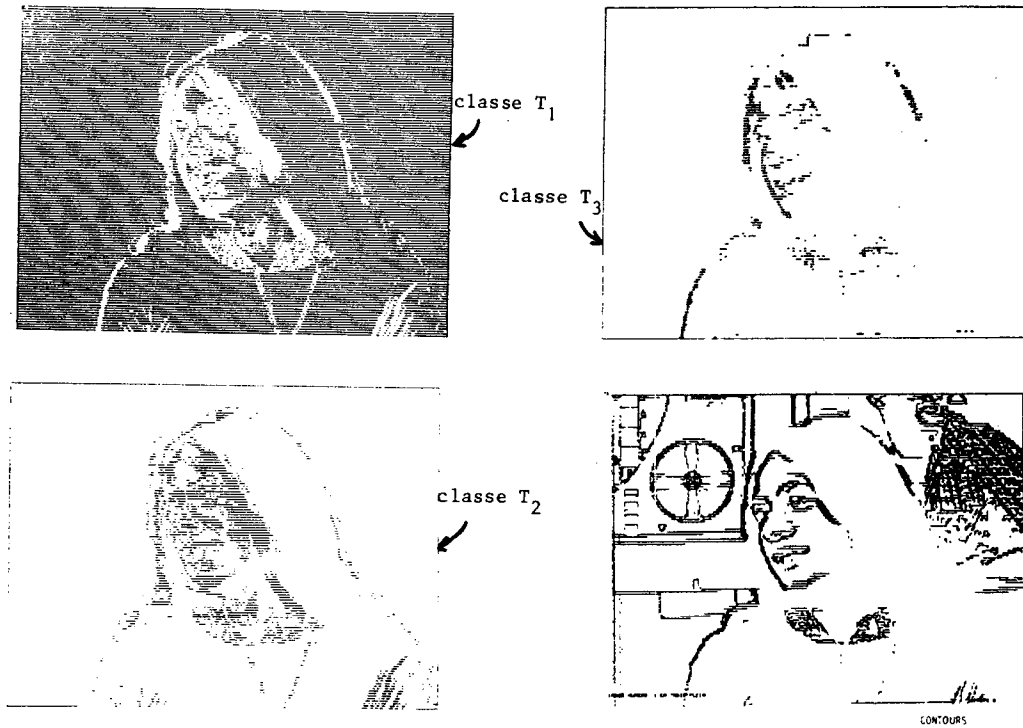
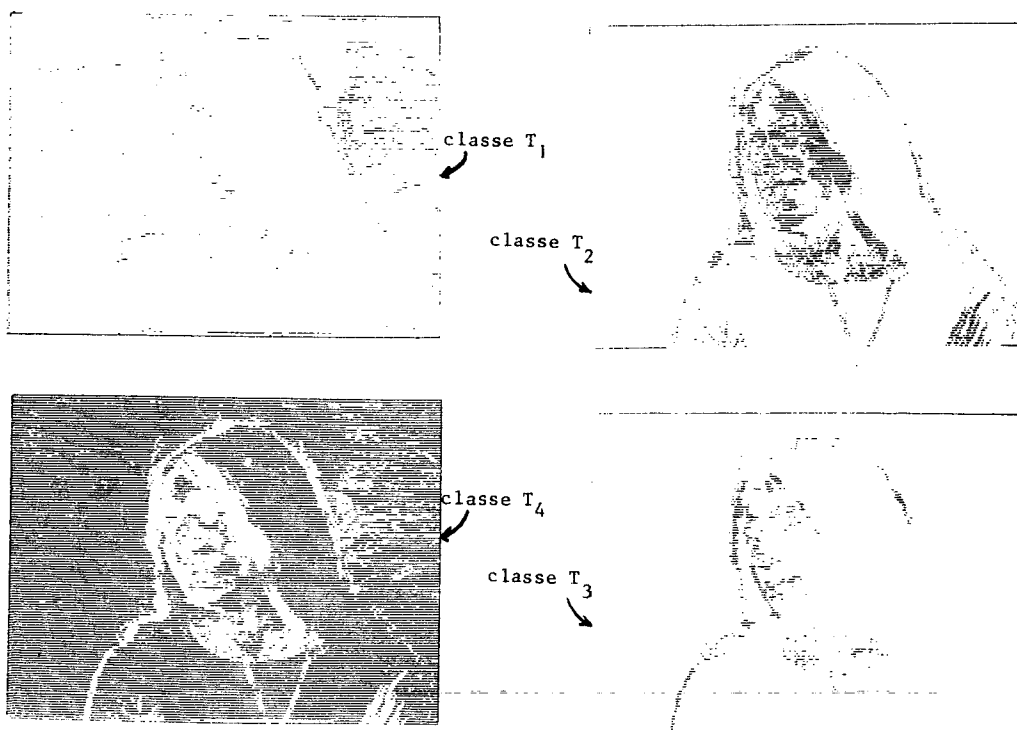


FIGURE B : Segmentation de l'Image "JUDITH" en 4 classes avant filtrage de la classe T₄



CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION .
 Adaptive multimode coding scheme for broadcast television sequences .

PHILIPPE ROBERT CLAUDE LABIT

FIGURE C : Partition de l'Image "JUDITH" en trois classes par segmentation avec filtrage.

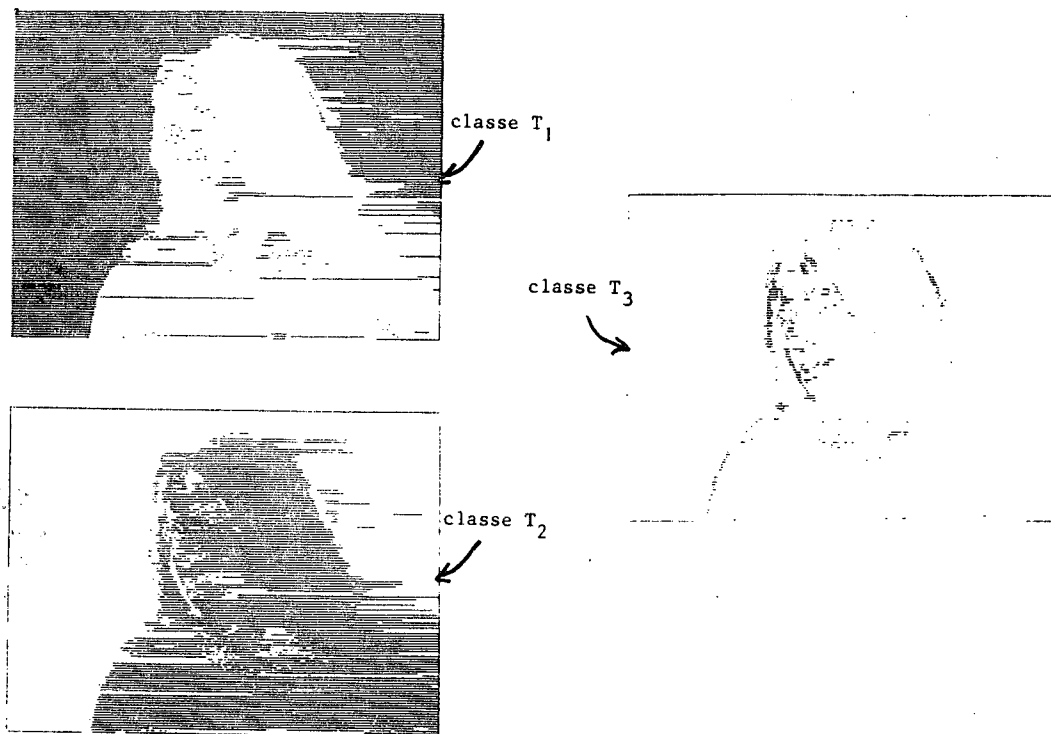
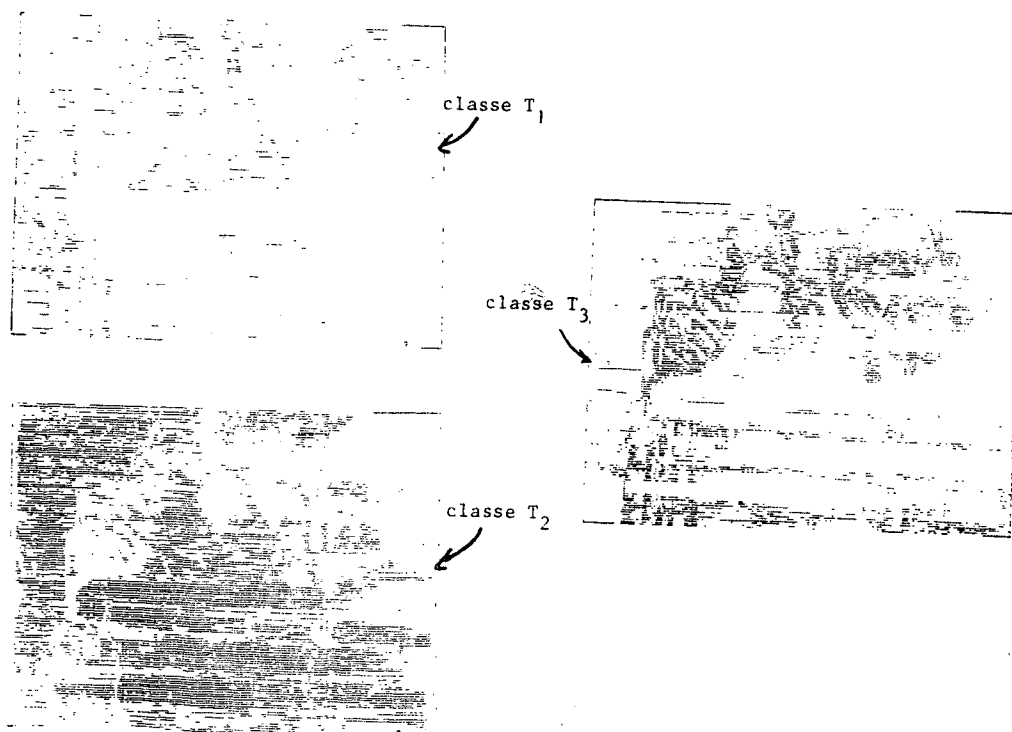


FIGURE D : Mêmes résultats sur l'Image "CHATEAU"



A N N E X E I V

**MULTIMODE CODER SCHEME AND OPTIMIZATION OF
ADDRESSING BIT RATE**

P. ROBERT, C. LABIT

PCS-RENNES 84

MULTIMODE CODER SCHEME AND OPTIMIZATION OF ADDRESSING BIT RATE

P. ROBERT C. LABIT

IRISA/Centre INRIA de RENNES

Av. du Général Leclerc

35042 RENNES Cedex FRANCE

PRELIMINARY REMARKS :

We have investigated sophisticated techniques involving edge detecting algorithms, motion estimation and multimode prediction strategies in order to reduce both spatial and temporal redundancies involved in real T.V sequences. We supported our approach with previous studies involving, on one hand, edge detection and following used to define intraframe adaptive coders ([1],[2]) and, on the other hand, motion estimation used to motion compensated coding ([3],[4]). The common framework involves, first a local level of observation (pixel level) both for feature estimation (edge orientation, motion) and for coding strategy, and also some overall processes: global filtering, divergence criteria and break tests on the estimators.

CONTRIBUTION :

Briefly, improving rate reduction involves two ways : reducing the number and the amplitude of prediction errors, and optimizing address coding if conditional coding techniques (and thus addresses) are used.

1) Prediction error bit rate

We don't focalise us on this topic (see [5],[6]). To optimally reduce both spatial and temporal redundancies, we use the following 3-modes coder (Fig.1):

- the mode 1 separates "fixed" from "moved" pixels as in usual conditional replenishment studies.

- The mode 2 separates motion compensated and uncompensated pixels using pel-recursive motion estimators [7] to which we add some improvements to enable and speed up the convergence of the estimation [4]

- The mode 3 provides us with adaptive spatial prediction [2].

2) Addressing bit rate [6] :

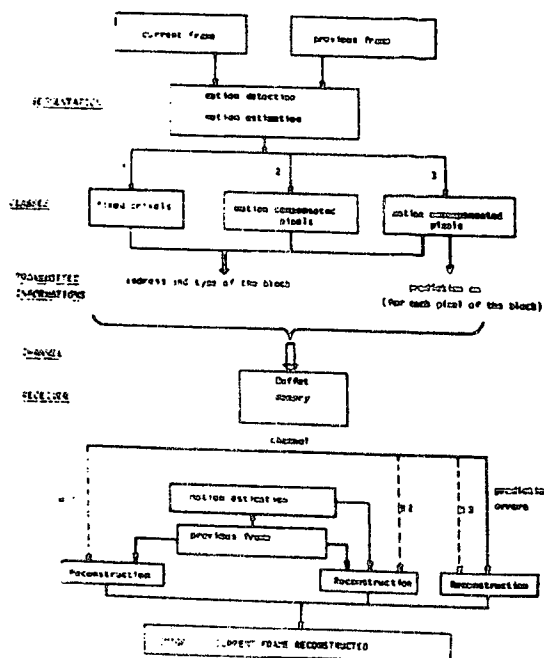
To reduce addressing bit rate, we have to decrease the number of clusters and code their addresses with optimal (1-D or 2-D) differential addressing coding. First a segmentation process classifies pixels under four headings (see Table 1) : fixed/moved pixels and motion compensated/uncompensated pixels. The ambiguity inherent in this segmentation -some pixels could be classified as fixed and motion compensated- enables a spatial hysteresis

process to give larger and homogeneous clusters.

Moreover, some efficient addressing techniques have already been used for fac-simile coding (EDIC, RAC, READ.) and may be helpful in our study. We extended them by using edge references correlated with prediction addresses.

CONCLUDING REMARKS

We are undertaking experiments on the sensitivity of these algorithms to different types of sequences, to transmission errors and have to confirm them with regard to visual tests.



SEGMENTATION :

heading ① : fixed and motion-uncompensated pixels
heading ② : displaced and motion compensated pixels

heading ④ : fixed and motion compensated pixels
heading ③ : others.

Class ④ is an additional but temporary subset because, all the pixels ④ will fall either under heading ① or ② as a result of the filtering process

Heading ④ $DFD1 < \lambda_1$ and $DFD1 > \lambda_1$
Heading ① $DFD1 < \lambda_1$ and $DFD1 > \lambda_1$
Heading ② $(DFD1 > \lambda_1 \text{ and } DFD1 < \lambda_1)$
or $(\lambda_1 < DFD1 < \lambda_2 \text{ and } DFD1 > DFD1)$
Heading ③ others....

HYSTERESIS :

The filtering process for pixels which fall under

heading ④ can be expressed as :

If X_{n-1} is heading H ① X_n is heading H ①
If X_{n-1} is heading H ② X_n is heading H ②
If X_{n-1} is heading H ③ X_n is heading H ③

Note: FD : frame difference

DFD : displaced frame difference

X_n and X_{n-1} : current and previous pixel

FIG. 1: Diagram of the coder
(from [6])

TABLE 1 (from [6])

- REFERENCES :
- [1]: C. RICHARD : IRISA Report n°124 (Nov. 79): "Estimation recursive de l'état local des contours et appl. à la prédiction ADPCM en T.V."
 - [2]: C. RICHARD, A. BENVENISTE, F. KRETZ : (to be published in IEEE Com. 84)
"Recursive estimation of local characteristics of edges in T.V pictures; as applied to ADPCM coding"
 - [3]: C. LABIT, A. BENVENISTE : NATO ASI Series (ed. by T.S Huang) Spr. Verlag 83
"Image sequence Processing and dynamic scene analysis : Motion estimation in T.V sequences."
 - [4]: C. LABIT, A. BENVENISTE : (This book)
"Edges and motions in T.V sequences"
 - [5]: P. ROBERT : IRISA Thesis (Nov 83)
"Définition d'un schéma de codage multimodes avec compensation de mouvement"
 - [6]: C. LABIT, P. ROBERT, J. DEWITTE : Int. Zurich Seminar on Digital Com. March 84
"Optimization of addressing bit rate in a multimode coder with motion compensation"
 - [7]: A.R. NETRAVALI, J.O. ROBBINS : BSTJ Vol 58 n°3 (March 79)
"Motion-compensated Television coding"

A N N E X E V

EDGES AND MOTIONS IN T.V. SEQUENCES APPLICATION TO INTRA/INTERFRAME DPCM CODING

C. LABIT, A. BENVENISTE

PCS RENNES 84

EDGES AND MOTIONS IN T.V SEQUENCES :

application to intra/interframe DPCM coding

C. LABIT A. BENVENISTE

IRISA/Centre INRIA de RENNES

AV. du General Leclerc

35042 RENNES Cedex FRANCE

Spatial edges and spatiotemporal edges (moving boundaries) correspond to stationarity breaks in the image sequence and thus, are the most critical loci for coding algorithms. It is quite important to observe the great correlation between the edge picture (given by spatial gradient detection process) and the frame difference picture (given by temporal thresholding on frame differences) (see Fig. 2 & 3) . In both cases, adaptive prediction techniques have to be used for improving predictability of these pixels. This analysis has been investigated by previous studies in spatial domain (edge detection and following [1],[2]) and temporal domain (edge motion estimation [3],[4]).

Briefly, we can summarize our approach into these stages: - a local estimation of edge features is first applied (orientation, motion and others state components).

- global processes are used to smooth these local observations. They are filtered recursively along the edges.

- At each pixel, a pel-recursive motion estimate is computed.

- However, using edge and motion models, model break tests are used to stop the filtering process when stationarity breaks occur and allow a fast reinitialization of the estimators.

The efficiency of such algorithms are described in [3],[4] and compared with several versions of per-recursive motion estimators. Fig. 4 shows the uncompensated pixels which remain.

To reduce optimally the number and amplitude of the prediction errors, we have investigated multimode scheme of coders combining intra and interframe predictors [5]. The "ideal" reduction that could be performed, is shown in Fig. 5 and 6 (respectively for the whole picture and for edge pixels only) when we keep the minimal prediction error (in absolute value) between three prediction strategies :

- a temporal prediction : $P(x,t)=I(x,t-1)$

- a temporal prediction with motion compensation : $P(x,t)=I(x-D,t-1)$.

- a spatial prediction : $P(x,t)=I(x-1,t)$. In [5], we used adaptive spatial predictors .

Performances are quite different according to the local characteristics of pixels : high contrast, high motion.... and this argues for adaptive and hybrid intra/interframe prediction schemes.

FIG 1

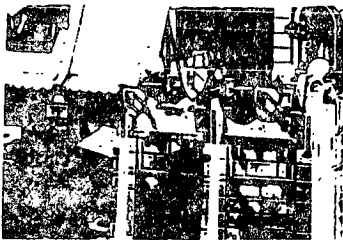


FIG 2

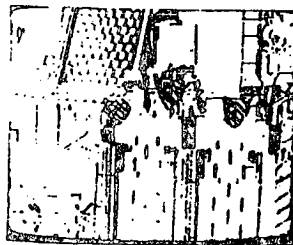


Fig. 1: input picture

Fig. 2: edge picture

Fig. 3: frame difference picture

Fig. 4: prediction error picture
(with motion compensation)

FIG 3

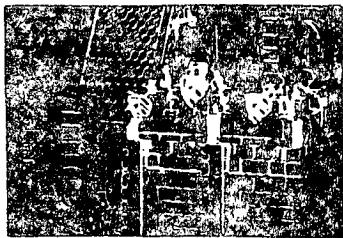
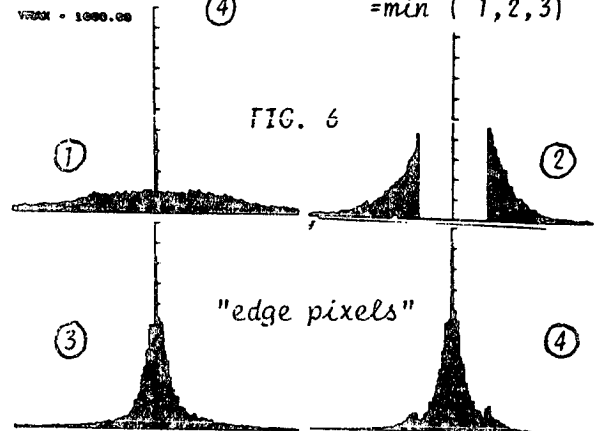
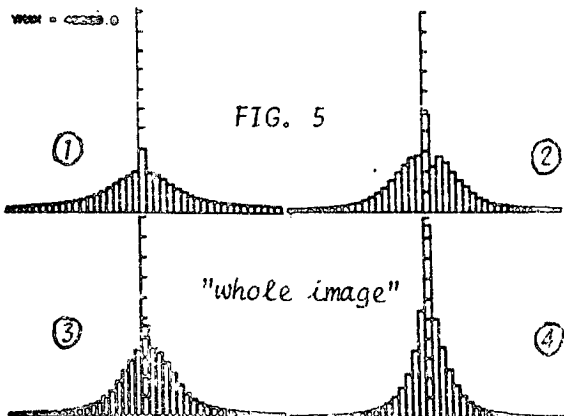


FIG 4



Fig. 5 and 6 : several prediction strategies

$$\begin{aligned} \textcircled{1} & P(x,t) = I(x,t-1) \\ \textcircled{2} & = I(x-1,t) \\ \textcircled{3} & = I(x-0,t-1) \\ \textcircled{4} & = \min(1,2,3) \end{aligned}$$



REFERENCES : [1] : F. KRETZ : Image sequence Processing and Dynamic scene analysis. Sp. Verlag 83 NATO Series

"Edges in visual scenes and sequences"; application to filtering, sampling and adaptive DPCM coding"

[2] : RICHARD C., BENVENISTE A., F. KRETZ : (to be published in IEEE Com 84)

"Recursive estimation of local characteristics of edges in T.V pictures"

[3] : C. LABIT, A. BENVENISTE : see 1

"Motion estimation in a sequence of T.V pictures"

[4] : C. LABIT, A. BENVENISTE : (submitted to IEEE Com)

"Contribution to edge-motion estimation for motion compensated coding in a sequence of T.V pictures".

[5] : P. ROBERT, C. LABIT : (this book)

"Multimode coder scheme and optimization of addressing bit rate"

A N N E X E . V I

**ESTIMATION DE PARAMETRES SPATIAUX (ORIENTATION ET MOUVEMENT)
DE CONTOURS D'UNE IMAGE EN VUE D'UN CODAGE ADAPTATIF : ETAT DE L'ART**

**3ÈME SYMPOSIUM FRANCO-SOVIÉTIQUE SUR LA T.V. NUMÉRIQUE
RENNES SEPTEMBRE 1984**

CLAUDE LABIT

(PAS D'ACTES)

A N N E X E V I I

**TEMPORAL PROPAGATION OF CHANNEL ERRORS IN
MULTITUDE ADAPTIVE CODERS**

C. LABIT, J.P. DE LA TRIBONNIERRE

**(SPIE CONFERENCE ON IMAGE CODING)
DEC 85, CANNES)**

TEMPORAL PROPAGATION OF CHANNEL ERRORS IN MULTIMODE ADAPTIVE CODERS

C. Labit*, J.P. de la Tribornière**

*IRISA/Centre INRIA de Rennes
35042 Rennes Cédex, France

**CCETT, Rue du Clos Courtel, BP 59
35510 Cesson Sévigné, France

Abstract

The aim of this study is temporal propagation of channel errors in coding schemes for Television Sequences.

Many studies have been performed to decrease spatial propagation in adaptive intraframe DPCM coder. All these methods could be extended to temporal domain. Briefly we distinguish, algorithmic robustness using efficient adaptive predictors, systematic protection with redundant information transmission, and the use of error correcting codes. This paper concerns more specifically the inherent robustness of coding algorithms.

A multimode intraframe/interframe coding algorithm using motion compensation is used and the locii where channel errors are introduced for simulation experiments are precisely defined : fixed areas, motion compensated areas, uncompensated areas. First experiments show that high degree of adaptivity in the sense of fast switching between prediction modes (intra/interframe) must decrease the error propagation in both temporal and spatial domains. Moreover, correcting improvements such as temporal refreshment of specific parameters, especially motion parameters, have been tested.

Introduction

The effect of channel errors in coding schemes has been mainly studied for adaptive intraframe coders : Many analysis and techniques that reduce propagation effects of these channel errors could be extended to interframe coders but also multimode (i.e. with several prediction domains) intra/interframe coders.

In view of the previous analysis in intraframe propagation, the approach taken in this paper is first to give some reasons of the potential propagations of channel errors, then to determine the experimental assumptions for the simulation, and finally to present some results and partial conclusions.

Analytical modeling of an adaptive coder subject to channel noise is difficult, more especially as the error propagation is tracked within an image sequence. So only heuristic descriptions are given here and statistical results try to summarize the phenomena of propagation and the effect of preliminary experiments of correcting methods.

Briefly, we distinguish the following correcting techniques :

- the inherent robustness of coding algorithms is the best way to bound error propagation without any increase of transmission bit rate
- some further parameters can be transmitted in addition to usual prediction errors (in DPCM coders). This method requires a higher bit rate but can expect the deletion of error propagation
- channel codes can be protected by using efficient (but redundant and expensive in terms of bit rate) error correcting codes.

This paper deals with the two first correcting techniques. First we recall some usual methods already applied to intraframe coders and describe how they can be extended to intra/interframe coders.

The next section shows precisely the multimode coder-decoder used for the simulation of a spatiotemporal communication system.

The two last sections specify the experimental assumptions and the results obtained by simulation of a noisy channel.

Algorithmic robustness

Some previous techniques have been developped to minimize and even suppress effects of error propagation with stable predictors. A more efficient robustness is achieved by using two-dimensional predictors and by introducing leak coefficients in the definition of predic-

tors or in the prediction function.

Predictors with leak coefficients

For the current pixel $x_{i,j}$ to be coded, the adaptive predictor $p(\hat{x}_{i,j})$ can be weighted by a leak coefficient μ which is equal, for instance, as proposed by Maxemchuk, Stuller and others, to the statistical mean of image luminances :

$$P_{i,j} = \alpha P(\hat{x}_{i,j}) + (1-\alpha) \mu$$

with

$$\mu = E(\hat{x}_{i,j})$$

Prediction functions with leak coefficients

The aim of this method is to introduce robustness within the adaptation of prediction function. This function is then defined by

$$P(\hat{x}_{i,j}) = \beta P(\hat{x}_{i,j}) + (1-\beta) \bar{P}(\hat{x}_{i,j})$$

with

$$P(\hat{x}_{i,j}), \bar{P}(\hat{x}_{i,j}) \text{ local adaptive predictor and fixed predictor of } x_{i,j}$$

and a weighting parameter β .

Concluding remark

If we combine these two techniques, the prediction of each pixel $x_{i,j}$ consists of three terms :

1. an adaptive "DPCM" term : $\alpha \beta P(\hat{x}_{i,j})$
2. a fixed "DPCM" term : $\alpha (1-\beta) \bar{P}(\hat{x}_{i,j})$
3. a fixed "PCM" term : $(1-\alpha) \mu$

The same improvements could be introduced in an interframe coding scheme with motion compensation. In this case, $P(\hat{x}_{i,j})$, $\bar{P}(\hat{x}_{i,j})$ and μ have to include temporal informations (motion estimate...).

Systematic replenishment of transmitted parameters

The effect of transmission errors can be reduced by sending redundant informations along the transmission channel. This technique increases the transmission bit rate but, however, enables efficient and fast breaks in divergence between coder and decoder. Adaptive algorithms need protected informations on which they can base. Thus we have tested the following protection :

- . to send PCM information within DPCM prediction errors. This is the usual technique for interframe conditional replenishment coders
- . to mix intraframe and interframe predictors
- . to send redundant informations along the channel. Several available parameters can be transmitted : the motion estimated vector in motion compensated coding schemes, discrete orientations of edges in adaptive intraframe coding, zonal addresses in run length coding, etc...

Note that some coding algorithms are provided for the systematic replenishment of these parameters and thus suppress causality constraints.

Multimode coding scheme

The background of this study is the previous implementation of multimode intra/interframe coding schemes applied to T.V. broadcast image sequences. Such a coder is thoroughly described in references 13, 14. Let us only recall its basic structure shown in Tableau 1.

The multimode coding scheme combines adaptive intraframe prediction, temporal conditional replenishment and interframe prediction with motion compensation.

Such a coding structure needs to transmit different kinds of information :

- . zonal addresses
- . prediction headings for each block
- . prediction errors only for motion uncompensated pixels.

The whole coding algorithm can be decomposed into the following steps :

Simulation context and experimental assumptions

Assumptions

We assumed that, only isolated errors appear, that the channel error rate is nearly 10^{-4} err/pix and that all the synchronization signals (line, field, frame synchronizations...) are efficiently protected.

The main consequence of this assumption is that the coder can be based on these signals and then, if the binary data flow of transmitted informations is desynchronized (i.e. if the zonal addresses or the number of prediction errors (on pixels ③) that had to be transmitted, are inconsistent, all these data of the current coded line are rejected and replaced by previously coded line using dissimulation techniques.

Simulation context

The experiments shown in this paper deals with :

- . the simulation of interframe dissimulation techniques
- . the observation of temporal propagation effects without any correcting technique
- . the simulations of some coarse correcting methods.

All these experiments use the image sequence "Voiture" which owns a lot of motion. They have been developed at CCETT on the sequence image processor "ARTIST". Some results have been only obtained on the central quarter of the image (see P2, P4, P5, P6 and P7) to minimize CPU time.

Error injection locations are first determined using a specific injection for each class of prediction (see P4). However it appears that the nature of the prediction mode where the channel errors are injected, exercises a little influence over the temporal effect of propagation. So random generated channel errors have been used to experiment the correcting methods (see P6, P7).

Experimental comments

The photographic results (P1 to P7) illustrate the series of experiments that we have performed.

Temporal dissimulation

Two dissimulation techniques are available according to the spatio-temporal causality constrain. An erroneous coded line can be replaced either by the previous line within the same frame or by the corresponding line in the previous frame (interframe dissimulation). Both these techniques give similar results because of equivalent correlations inherent in spatial or spatio-temporal neighbourhoods. We only present results on temporal dissimulation. The picture P2 shows the binary image of thresholded differences in absolute values between decoded pictures with and without temporal dissimulation. It appears that the temporal dissimulation creates homogeneously granular noise which is however not critical visually (see P3). So the feasibility of temporal dissimulation is proved using such an adaptive multimode coder.

Coarse correcting methods

This study is still in progress and so only coarse correcting methods have been tested.

The temporal effects of a channel-error propagation within the image sequence is shown (P5 at picture (t_0+1) and P7 at picture (t_0+17)). We can see quite severe divergence between reconstructed images respectively at the coder and decoder. This divergence has an asymptotic behaviour (around picture (t_0+15)) and after some frames, we observe a nearly constant distortion.

Two coarse correcting methods are presented here :

- . a systematic replenishment of four coded lines is performed. This method does not create great changes (see P6) and has to be improved
- . however, when motion vectors are additionally transmitted (at the frequency of one vector for 10 pixels) and are assumed to be right, the divergence between the coder and the decoder is well attenuated. The photograph P7 shows the residual distortion observed at time (t_0+17) when an error injection with a $4 \cdot 10^{-4}$ error channel rate is applied to several successive picture $(t_0$ to $t_0+7)$.

Conclusion

The aim of this study was temporal propagation of channel errors in coding schemes which have high degree of adaptivity : the design of multiple prediction modes in both spatial and temporal domains and a fast switching between them authorizes a more efficient algorithmic robustness towards (temporal) propagation. As we have seen by temporal dissimulation experiments, channel error injections introduce granular noise that could be not too critical visually in magnitude. Using coarse correcting techniques, it appears that temporal propagation effects can be decreased and even deleted. However a more realistic trade-off between an available correcting technique and a reasonable cost of additional transmission bit rate (nearly 5% of the global bit rate) will have to be defined in the further experiments.

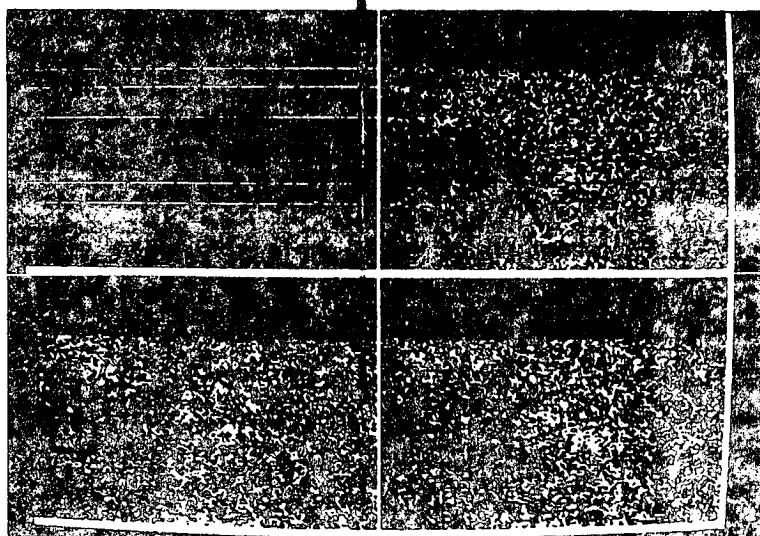
References

- 1 R.J. Arguello, H.R. Sellner, J.A. Stuller : The effects of Channel errors in the differential pulse-code-modulation. Transmission of sampled Imagery. IEEE Trans. on Com., vol.19, n°6, Déc. 71.
- 2 D.J. Connor : Techniques for reducing the visibility of Transmission errors indigitally encoded video signals. IEEE Trans. on Com., vol.21, n°6, June 1973.
- 3 M.C.W. Van Buul : Hybrid D-PCM. A combination of PCM and DPCM. IEEE Trans. 78.
- 4 V. Guionnet, P. Sallio : Système expérimental de T.V. numérique à 52Mb/s. Evaluation subjective de l'effet des erreurs de transmission. Revue de Radiodiffusion-Télévision, n°54, 1978.
- 5 R. Steele, D.J. Goodman, C.A. McGonegal : A difference detection and correction scheme for combating DPCM transmission errors. IEEE Trans. on Com., vol.27, n°1, Janvier 1979.
- 6 T. Kremers, M.C.W. Van Buul : Hybrid D-PCM for joint source/channel encoding. Philip Report, 1979.
- 7 N.F. Maxemshuk, J.A. Stuller : Reduction of Transmission Error propagation in adaptively predicted DPCM encoded pictures. BSTJ, vol.58, n°6, Juillet-Août 1979.
- 8 D. Barba, J. Dewitte : Reduction of effects of source noise and transmission errors in adaptive DPCM. PCS 81, Montréal.
- 9 H. Yamaguchi, Y. Hatori, H. Yamamoto : Transmission error propagation in 2-D adaptive DPCM Prediction. PCS 81 Montréal.
- 10 K.N. Ngan, R. Steele : Enhancement of PCM and DPCM Images corrupted by transmission errors. IEEE Trans. on Com., vol.30, n°1, Janvier 1982.
- 11 F. Kretz : Codage MICD à prédiction adaptative en T.V. : qualité visuelle, réduction de débit et susceptibilité aux erreurs de transmission. Annales des Télécom., tome 37, n°7, 8, Juillet-Août 1982.
- 12 W.W. Peterson, E.J. Weldon : Error correcting codes. MIT Press, 1972.

- 13 P. Robert : Définition d'un schéma de codage multimodes, avec compensation de mouvement pour les séquences d'images de télévision. Thèse Docteur-Ingénieur, Université de Rennes I, Novembre 1983.
- 14 C. Labit, P. Robert, J. Dewitte : Optimization of addressing bit rate in a multimode coder with motion compensation for broadcast television. International Zurich Seminar, Fev. 1984.

P1

Original image and location of dissimulated lines in overlay mode.



P2

Thresholded ($\lambda=6$) differences between decoded pictures with and without temporal dissimulation

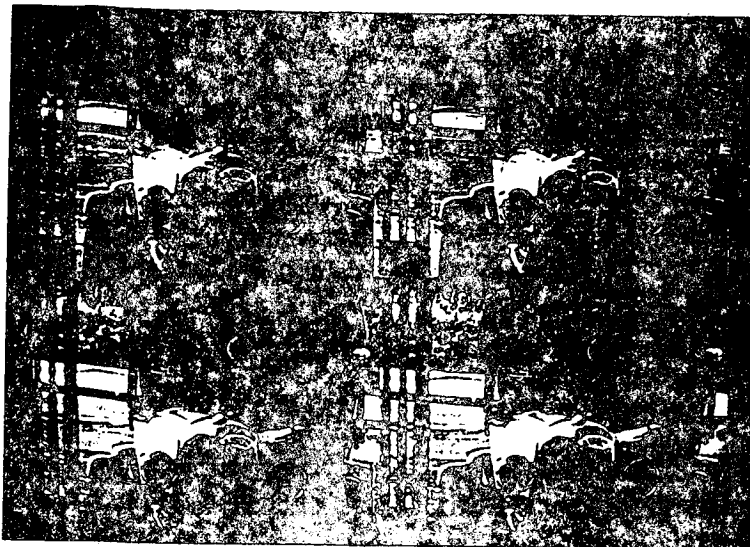
image t_0^*	image t_0+1
image t_0+2	image t_0+3

$*t_0$ is the picture number where the temporal dissimulation starts up.

P3

Reconstructed image (t_0+1) at the decoder after temporal dissimulation processing.





p4

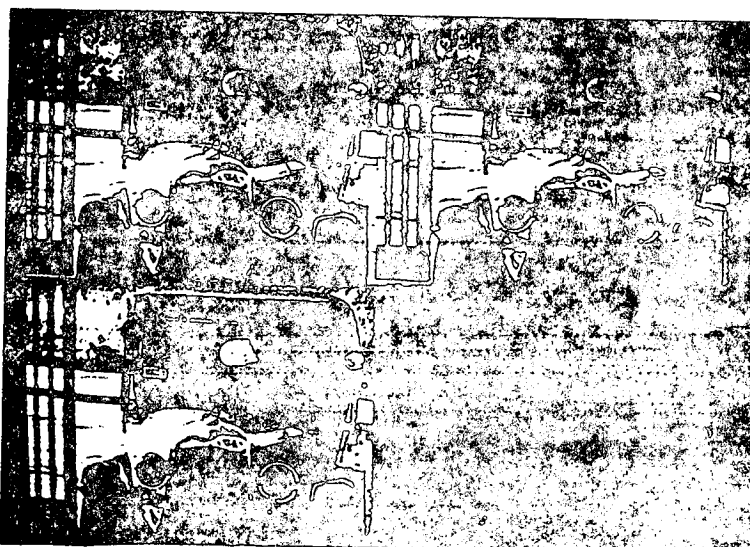
Part (1/4) of the original image and location of error injections at picture (t_0)

class 1	class 2
class 3	} class i

P5

Temporal propagation at picture (t_0+1)

original picture	reconstructed image at the coder
	↓
reconstructed image at the decoder	thresholded ($\lambda = 6$) differences



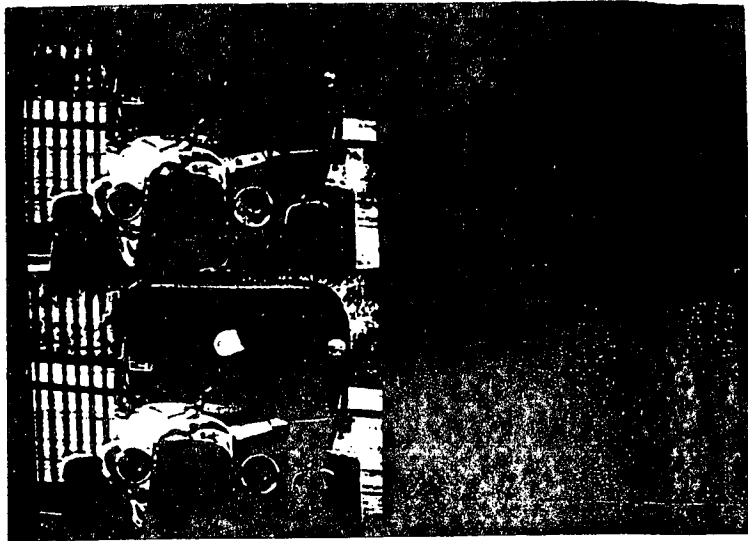
P6

id as P.5 with a systematic replenishment of four coded lines

P7

error injection ($4 \cdot 10^{-4}$ rate) at picture t_0 . Observation of temporal distortions \rightarrow at picture t_1

error injection ($4 \cdot 10^{-4}$) at pictures t_0, \dots, t_0+7 . Observation of temporal distortions at picture t_1 using motion-estimates replenishment.



↑
decoded pictures
 $t_1 = (t_0+17)$

↑
thresholded ($\lambda=6$) differences
between reconstructed images
at the coder and the decoder.

A N N E X E V I I I

**TEMPORAL EFFECTS OF CHANNEL ERRORS
IN INTRA/INTERFRAME MULTIMODE CODERS**

J.P. DE LA TRIBONNIERE, C. LABIT

PCS - Tokyo 86

TEMPORAL EFFECTS OF CHANNEL ERRORS IN INTRA/INTERFRAME MULTIMODE CODERS

Jean-Pierre de la TRIBONNIERE*, Claude LABIT**

*CCETT, Rue du Clos Courtel, BP 59 - 35510 Cesson-Sévigné
Tél.:(99)02.41.11

**IRISA/Centre INRIA de Rennes - 35042 Rennes Cédex
Tél.:(99)36.20.00 - Télex: Unirisa 950473F

INTRODUCTION

The effect of channel errors in coding schemes has been mainly studied for adaptive intraframe coders ([1], [4], [5]). Many analysis and techniques that reduce propagation effects of channel errors could be extended to interframe coders and even to multimode (i.e. with several hybrid prediction domains) inter/intraframe coders.

Analytical modeling of an adaptive coder subject to channel noise is difficult, more especially if the error propagation is tracked within an image sequence. So only heuristic descriptions are given here and statistical results will try to summarize the phenomena of propagation and the effect of correcting methods.

Briefly, we could implement the following correcting techniques :

- *the inherent robustness of coding algorithms is efficient to bound error propagation without any increase of the transmission bit rate,
- *some further parameters can be transmitted in addition to usual prediction errors (in DPCM coders). This method requires higher bit rate but can expect the deletion of error propagations,
- *channel codes can be protected by using efficient (however, redundant and expensive in terms of bit rate) error correcting codes [6, 9].

This paper deals with the two first correcting techniques.

MULTIMODE CODING SCHEME

The background of this study is the previous implementation of multimode intra/interframe coding schemes applied to T.V. broadcast image sequences. Such a coder is thoroughly described in references [10, 11]. Let us only remind its basic structure shown in Tab. I.

The multimode coding scheme combines adaptive intraframe prediction, temporal conditional replenishment and interframe prediction with motion compensation.

EXPERIMENTAL ASSUMPTIONS

We assumed that, only isolated errors appear, that the channel error rate is nearly 10^{-4} err/pix and that all the synchronization signals are efficiently protected. Thus the coder can be based on these signals and then, if the binary data flow of transmitted informations is desynchronized, all these data of the current coded line are rejected and replaced using dissimulation techniques.

EXPERIMENTS AND PRELIMINARY CONCLUSIONS

Simulation of two methods of dissimulation have been developed. The erroneous transmitted line is replaced by either the corresponding line in the previous frame (interframe dissimulation) either the previous line in the current coded picture (intra-frame dissimulation). Both techniques give equivalent results in terms of objective distortion. These distortions are not visually critical.

We have simulated different locii for error intrusion, especially within each class of prediction (i.e. on fixed pixels, motion-compensated pixels and others...) and also generated random injection.

We observe an asymptotic behaviour with regard to the divergence between coder and decoder. In order to measure objective distortion, we have drawn curves of thresholded differences in absolute values between original sequences and reconstructed sequences after an erroneous transmission. Visual effects of temporal propagation will be shown further during the session.

1. R.J. Arguello, H.R. Settler, J.A. Stuller: "The effects of Channel errors in the differential pulse-code-modulation Transmission of sampled Imagery". IEEE Trans. on Com., vol.19, n°6, Déc. 71.
2. D.J. Connor: "Techniques for reducing the visibility of Transmission errors digitally encoded video signals". IEEE Trans. on Com., vol.21, n°6, June 73.
3. M.C.W. Van Brui: "Hybrid D-PCM: A combination of PCM and DPCM". IEEE Trans. 78.
4. A. Guionnet, P. Sallio: "Système expérimental de T.V. numérique à 52Mb/s. Evaluation subjective de l'effet des erreurs de transmission". Revue de radio-diffusion-Télévision, n°34, 1978.
5. N.F. Moysenbush, J.A. Stuller: "Reduction of Transmission Error propagation in adaptively predicted DPCM encoded pictures". BSTJ, vol.58, n°6, Juillet-Aout 1979.
6. P. Stammitz: "Error protection of 34Mb/s DPCM encoded TV signal with multiple error correcting BCH codes". International Zurich Seminar on Digital Communications, Mars 1980.
7. U. Barot, J. Dewitte: "Reduction of effects of source noise and transmission errors in adaptive DPCM". PCS 81, Montréal.
8. F. Kretz: "Codage MCD à prédiction adaptative en T.V.: qualité visuelle, réduction de débit et susceptibilité aux erreurs de transmission". Annales des Télécom., tome 37, n°7-8, Juillet-Aout 1982.
9. W.H. Peterson, E.J. Weldon: "Error correcting codes". MIT Press, 1972.
10. P. Robert: "Définition d'un schéma de codage multimodes, avec compensation de mouvement, pour les séquences d'images de télévision". Thèse Docteur Ingénieur, Université de Rennes I, Nov. 1983.
11. P. Robert, P. Robert, J. Dewitte: "Optimization of addressing bit rate in a multimode coder with motion compensation for broadcast television". International Zurich Seminar, Feb. 1984.

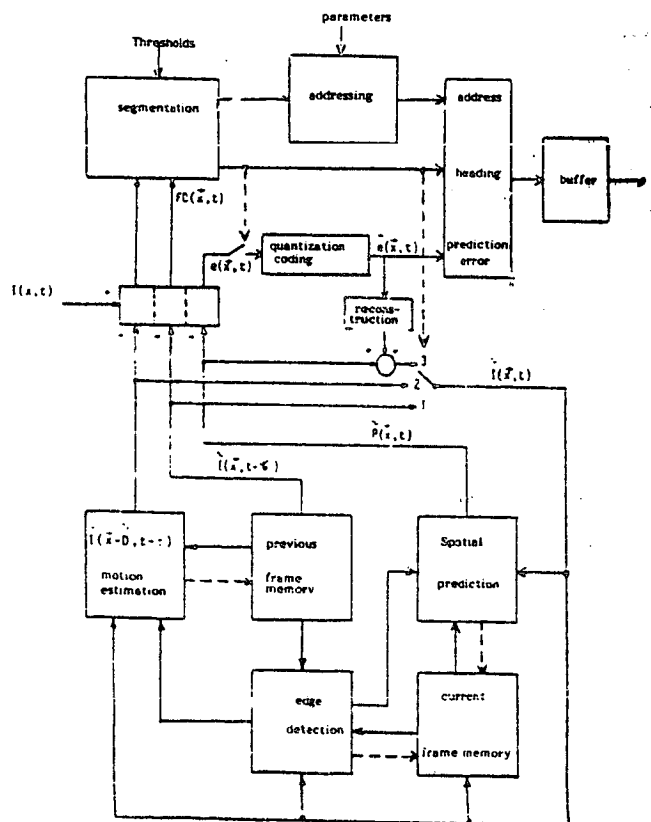


Tableau 1: Multimode coder scheme

A N N E X E I X

**DISSIMULATION ET CORRECTION DE LA
PROPAGATION TEMPORELLE D'ERREURS DE TRANSMISSION**

J.P. DE LA TRIBONNIERE, C. LABIT

Deuxième colloque image / Second Image Symposium

traitement, synthèse, technologie et applications / image processing computer generated images, technology and applications

Nice, Avril 1986

DISSIMULATION ET CORRECTION
DE LA PROPAGATION TEMPORELLE D'ERREURS DE TRANSMISSION
DISSIMULATION AND CORRECTING TECHNIQUES
TO TEMPORAL EFFECTS OF CHANNEL ERRORS

J.P. de la Tribonnière* - C. Labit**

*CCETT, Rue du Clos Courtel, 35510 Cesson-Sévigné - **INRIA de Rennes, Beaulieu, 35042 Rennes Qx

RESUME

Les systèmes adaptatifs utilisés en codage d'images nécessitent une forte protection vis-à-vis des erreurs sur les informations transmises qui peuvent venir perturber un canal de transmission. De nombreuses techniques de protection d'algorithmes ont été essayées dans le domaine du codage intrimage. Nous étendons l'étude au cas de codeur adaptatif multimodes intra et interimage avec compensation de mouvement. Dans ce cas l'introduction et l'étude de la propagation des erreurs peut se faire à différents niveaux. Nous détaillons dans ce papier ces différents niveaux et présentons les simulations réalisées pour l'étude de techniques de correction et dissimulation d'erreurs.

Ces expériences montrent qu'une structure fortement adaptative du codeur empêche une divergence trop importante entre le codeur et le décodeur. Cependant, aucun effet d'"atténuation" temporelle n'est observé si aucune technique de correction n'est employée. Le rafraichissement "régulier" des paramètres de mouvement est apparu comme l'une des techniques de correction les plus efficaces pour combattre la propagation temporelle des erreurs dans le cadre de schémas de codage comportant un mode de prédiction par compensation de mouvement.

SUMMARY

The effect of channel errors in coding schemes has been mainly studied for adaptive intraframe coders. These analysis and techniques that reduce propagation effects of channel errors could be studied and extended to interframe coders and even to multimode (i.e. with several hybrid prediction modes, inter/intraframe coders).

We present some experiments about dissimulation, error injection and correcting techniques applied to the different levels of the coder scheme.

These experiments show that high degree of adaptivity prevents the reconstructed image from a tremendous divergence. However, no temporal vanishing effect occurs if correcting techniques are not implemented. Correcting improvements such as temporal refreshment of specific parameter, especially motion parameter, have been tested and thus enable an effective decrease of the temporal propagation of channel errors in the case of coding scheme using motion compensated predictors.

DISSIMULATION ET CORRECTION DE LA PROPAGATION TEMPORELLE
D'ERREURS DE TRANSMISSION - JP de la Tribonnière, C.Labit
DISSIMULATION AND CORRECTING TECHNIQUES TO TEMPORAL EFFECTS OF
CHANNEL ERRORS - JP de la Tribonnière, C.Labit

I. Introduction : Bref aperçu des techniques utilisées

De nombreuses études ont porté sur la propagation des erreurs de transmission dans le cadre de schémas de codage adaptatif ADPCM intraimage [9]. Les notions introduites et protections apportées peuvent être reprises et étendues au cas de propagation temporelle interimage qui constitue la partie délicate et peu étudiée des propagations d'erreurs dans le cadre de schémas de codage adaptatif interimage. La revue bibliographique que nous avons effectuée et dont quelques références sont données en Annexe, ne se veut pas exhaustive mais fournit l'ensemble des outils et notions élémentaires expliquant la méthodologie adaptée ainsi que les expérimentations proposées.

Cette étude est menée en collaboration entre le CCETT et l'IRISA/Centre INRIA de Rennes.

Plusieurs types de protection, de nature fortement distincte peuvent être apportés aux schémas de codage adaptatif. Afin de réduire les effets de propagation des erreurs de transmission, distinguons schématiquement :

- la protection par robustesse algorithmique,
- la protection "systématique",
- la protection par codes redondants.

Le dernier aspect concerne les études de codage-canal optimal qui pour un débit fixé autorisent l'introduction de bits de "redondance" dans les mots de code afin de détection/corriger une ou plusieurs erreurs. Des études complètes et comparatives ont été menées [10,11] (codes cycliques, BCH...) dans le cadre du codage d'images de T.V. Ce type de protection ne sera pas l'objet de notre étude.

I.1. - Protection algorithmique [5]

Cette protection vise à la conception d'algorithmes de codage intrinsèquement robustes vis-à-vis des erreurs de transmission. Cette robustesse peut être obtenue par l'utilisation de prédicteurs bidimensionnels stables (plus robustes que les prédicteurs 1-D) et par l'introduction de coefficients de fuites ("leak coefficients") dans la définition des prédicteurs ou dans celle de la fonction de prédiction.

Ces techniques ont d'ores et déjà été testées et comparées dans le cas de codeurs intraimages ou interimages simples. Elles méritent d'être étendues au cas interimage avec compensation de mouvement.

I.2. - Protection "systématique" [3]

Il s'agit dans ce cas de protéger les informations essentielles à transmettre par des informations fortement redondantes, faciles à protéger ou par des informations annexes permettant une reconstruction sûre du signal d'images en cas d'introduction d'erreurs sur le canal. Citons les quelques exemples suivants :

. *Envoi d'informations PCM* : dans un schéma de codage DPCM. Cet envoi se fera de manière

cyclique assujettie à la contrainte d'un débit fixé. La protection des codeurs interimages à rafraîchissement conditionnel peut ainsi s'effectuer par l'envoi cyclique d'informations PCM spatiales d'images parmi les erreurs de prédiction interimages.

. *Envoi d'informations annexes* : ex: vecteurs de mouvement, orientations discrètes de contour, adresses de zones. Les informations annexes permettent de définir quel prédicteur est choisi dans le cadre de la fonction de prédiction adaptative et interviennent bien souvent dans le calcul lui-même du prédicteur

Exemple :

le vecteur \vec{D} pour les schémas de codage à compensation de mouvement.

le paramètre θ d'orientation des contours pour les schémas de codage, intraimages DPCM avec prise en compte des contours.

Cet envoi d'informations annexes permet de stopper la divergence et de recalculer la fonction adaptative de prédiction. Certains schémas de codage prévoient l'envoi systématique de ces informations afin de ne pas s'astreindre à la contrainte de causalité.

I.3. - Protection-canal par codes redondants [10]

Celle-ci a fait l'objet d'études comparatives et de tests de codes redondants (codes BCH) pour la transmission du signal TV à 34 Mb/s.

Ces codes redondants permettent de détecter les erreurs et de corriger les erreurs ($p' \leq p$) dans un mot binaire de longueur n . L'efficacité de ces codes est mesurée par le rapport

$$\frac{\text{nb de bits d'info. utiles}}{\text{nb de bits total}} = \frac{n-k}{n} \quad (k \text{ bits de redondance})$$

Ces techniques de protection sont à considérer comme protection-canal à appliquer de manière systématique ou sélective suivant l'information à protéger.

II. Cadre de l'étude

Nos études précédentes en codage ont porté sur les schémas adaptatifs intraimages et hybrides intra/interimages avec compensation de mouvement.

Le schéma de codage envisagé dans le cadre de cette étude est celui détaillé dans [13,14] et résumé par le schéma-bloc suivant.

Cette structure de codeur prévoit donc l'envoi sur le canal de transmission de trois types d'informations :

1. des informations d'adresses : le codage se fait par plages à stratégie de prédiction constante. Ces adresses peuvent être monodimensionnelles ou bidimensionnelles et donc ce dernier cas, avec ou sans prédiction d'adresses.

2. des informations de type de prédiction :

DISSIMULATION ET CORRECTION DE LA PROPAGATION TEMPORELLE
D'ERREURS DE TRANSMISSION - JP de la Tribonnière - C.Labit

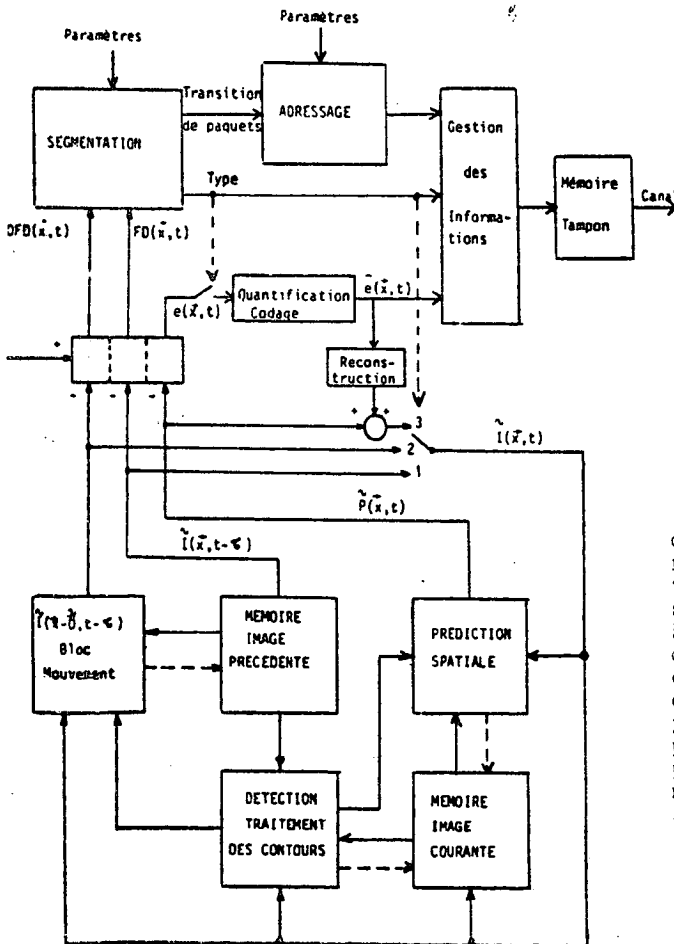
DISSIMULATION AND CORRECTING TECHNIQUES TO TEMPORAL EFFECTS OF
CHANNEL ERRORS - JP de la Tribonnière - C.Labit

trois stratégies de prédiction ont été retenues :

- une prédiction interimage "simple"
- une prédiction interimage avec compensation de mouvement
- une prédiction intrainage (" $(A+C)/2$ ")

3. des informations d'erreurs différentielles de prédiction. Celles-ci sont intrainages et correspondent aux points de l'image non compensés en mouvement.

L'influence et la propagation des erreurs de transmission va dépendre de la nature de l'information (adresse, type, erreur,...) où est injectée l'erreur. Bien que l'ensemble des effets s'enchevêtre, pour des raisons de clarté, nous supposons que l'erreur est introduite sélectivement sur l'un ou l'autre des types distincts d'informations.



SCHEMA GENERAL DU CODEUR

III. Analyse du problème

III.1. Cause d'erreurs dans les schémas ADPCM

Nous reprenons, en l'étendant au cas de codeurs multimodes intra/interimages, la classification donnée dans [7]. Bien que l'ensemble des causes et effets de la propagation d'erreurs se confond au niveau de l'image reconstruite finale, il est intéressant de cerner quels sont les niveaux d'introduction et de propagation d'erreurs dans de tels schémas de codage.

a - erreur dans la prédiction

La prédiction (et donc l'erreur de prédiction calculée corollairement) peut se trouver erronée sous l'effet conjugué ou non de trois mécanismes :

erreur dans le calcul lui-même de la prédiction : les pixels reconstruits intervenant dans le calcul de prédicteur peuvent avoir été erronés précédemment.

erreur dans la stratégie de prédiction : dans le cas d'une politique adaptative de prédiction ou de quantification, les tests nécessaires au choix du prédicteur ou du quantificateur peuvent conduire à un choix erroné s'ils portent sur des pixels ou informations précédemment erronés.

erreur dans le calcul de paramètres annexes : dans l'exemple du schéma de codage multimodes que nous proposons, ces paramètres annexes intervenant dans le calcul des prédicteurs sont :

- l'estimation de l'orientation discrète des contours
- l'estimation du mouvement \vec{D} en chaque point.

b - erreur dans l'adressage

Si le schéma de codage choisi est un codeur non systématique (c'est-à-dire où une erreur de prédiction n'est pas envoyée systématiquement en chaque pixel) alors il est nécessaire de transmettre les adresses des blocs à stratégies de prédiction distincte. L'introduction d'erreurs à ce niveau crée des blocs complets de pixels erronés. Si de plus un adressage différentiel (1-D ou 2-D) est employé, la synchronisation du train binaire est perdue dramatiquement. Cette cause d'erreurs ne peut être combattue par des méthodes systématiques de rafraîchissement ou dissimulation si l'erreur est détectée.

c - erreur dans le type

Une erreur dans le type de prédiction conduit le décodeur à rechercher le prédicteur du point d'image dans une classe de prédicteurs différente de celle choisie en amont au codage. Un point fixe peut être déclaré en mouvement compensé ou vice-versa. La synchronisation du train binaire peut être également perdue s'il y a interversion entre un type avec envoi d'erreurs de prédiction ou sans.

DISSIMULATION ET CORRECTION DE LA PROPAGATION TEMPORELLE
D'ERREURS DE TRANSMISSION - JP de la Tribonnière, C. Labit

DISSIMULATION AND CORRECTING TECHNIQUES TO TEMPORAL EFFECTS OF
CHANNEL ERRORS - JP de la Tribonnière, C. Labit

III.2. -- Injection des erreurs

Nous distinguerons trois niveaux d'observations :

a - injection des erreurs

Les études précédentes concernant les erreurs de transmission [9] ont souvent introduits de manière sélective le(s) erreur(s) :

- dans les zones homogènes
- dans les zones contours

puisque à ces deux classes correspondaient les deux stratégies de prédiction dans un schéma ADPCM intraimage. Pour le domaine interimage, il conviendra de même d'introduire sélectivement les erreurs :

- dans les zones fixes
- dans les zones en mouvement
- sur les contours d'occlusion en mouvement (zones critiques pour les systèmes de codage à compensation de mouvement).

b - observation immédiate au point ou ligne courante

Cette observation concerne l'étude de la propagation de l'erreur localement (voisinage monodimensionnel) autour du point courant où l'erreur est introduite.

c - propagation bi ou tri-dimensionnelle

Il s'agit de la propagation spatiale et temporelle des erreurs ; bien souvent les diverses causes de propagation de ces erreurs dans de telles zones se combinent et deviennent complexes à analyser.

Deux niveaux d'études sont à distinguer :

- propagation des erreurs sur les prédictions
- propagation des erreurs sur les schémas de prédiction.

IV - Expérimentations et analyses

IV.1. -- Hypothèses expérimentales

Nous supposerons les hypothèses suivantes concernant le type d'erreurs introduites et le train binaire transmis :

- . les erreurs se présentent de manière isolée et non en paquets
- . le taux d'erreurs choisi pour ces expérimentations restera compris entre $2.8 \cdot 10^{-6}$ à $4 \cdot 10^{-4}$ ce qui correspond à un minimum d'une erreur injectée par image et à un maximum de 80 erreurs par trame. Ces bornes ont été choisies pour simuler de manière réaliste le fonctionnement d'une ligne de transmission courante.

- . les synchronisations Image, Trame, ligne sont supposées fortement protégées. Le codeur est supposé posséder la faculté d'autodétection d'erreurs de synchronisation du train binaire ; en cas de détection d'erreurs dans la cohérence de ce train, il y a rejet du train-ligne et remplacement par des techniques de dissimulation que nous présentons ci-dessous.

Les simulations ont porté sur la séquence normalisée COST : VOITURE. La fréquence d'échantillonnage des luminances est de 13,5 Mhz ; des suites temporelles de 10 à 50 images ont été utilisées pour tester les effets de propagation temporelle des erreurs de transmission. De plus, afin de minimiser le temps-calcul de chaque traitement, les expérimentations auront lieu bien souvent sur le quart de l'image. (Voir images en annexe).

IV.2. -- Techniques de dissimulation

Les techniques de dissimulation classiquement utilisées travaillent dans le domaine intraimage [9] mais doivent respecter les contraintes de causalité dans le cas où des prédicteurs bidimensionnels spatiaux sont utilisés. Dans le domaine interimage, une dissimulation temporelle peut être envisagée et méritera d'être testée avec toutefois la contrainte similaire de causalité temporelle.

Nous avons testé deux types de dissimulation :

- . la dissimulation de la ligne courante déclarée erronée par la ligne précédente de la même image (dissimulation intraimage)

- . la dissimulation de la ligne courante déclarée erronée par la même ligne dans l'image précédente (dissimulation interimage).

On aurait pu penser à une différence de performances importante dans le cas de configurations d'images spécifiques (fort mouvement + avantage à la dissimulation intraimage, contours horizontaux + avantage à la dissimulation interimage).

Toutefois ces configurations sont locales et ne peuvent donc rendre les performances de telle ou telle dissimulation meilleures de manière globale sur l'ensemble de la ligne dissimulée.

Les résultats donnés en Annexe concernent une illustration de la localisation des lignes dissimulées et de la propagation temporelle des erreurs introduites par ces techniques de dissimulation. Nous présentons également des résultats comparatifs de ces deux techniques de dissimulation : visuellement, celles-ci ont été jugées équivalentes et seront donc à employer suivant la commodité pour le codeur.

IV.3. -- Injection des erreurs

Nous avons détaillé au paragraphe III la méthodologie d'injection des erreurs aux différents niveaux d'un codeur ADPCM intra/interimage.

En reprenant les notations suivantes :

- zones 1 : zones d'image à prédiction interimage "simple"
- zones 2 : zones d'images à prédiction interimage avec compensation de mouvement
- zones 3 : les autres ...

Dans un premier temps, nous avons cher-

DISSIMULATION ET CORRECTION DE LA PROPAGATION TEMPORELLE
D'ERREURS DE TRANSMISSION - JP de la Tribonnière, C. Labit

DISSIMULATION AND CORRECTING TECHNIQUES TO TEMPORAL EFFECTS OF
CHANNEL ERRORS - JP de la Tribonnière, C. Labit

ché à injecter les erreurs de transmission d'une manière sélective sur les différents types d'informations transmises par le codeur :

a - erreurs de transmission sur les erreurs de prédiction

Ces erreurs injectées concernent donc exclusivement les zones 3 qui représentent principalement les contours d'occlusion en mouvement. Ce type d'erreurs de transmission n'affecte en rien la synchronisation du train binaire mais erronne les prédictions associées.

b - erreurs de transmission sur les adresses

Pour combattre ce type d'erreurs, seuls les adressages différentiels monodimensionnels apparaissent suffisamment réalistes. Le gain en débit apporté par des schémas plus complexes d'adressage différentiels bidimensionnels ne semble pas suffisant pour autoriser une déficience importante en robustesse vis-à-vis des erreurs de transmission. Dans le cas d'adressage différentiel 1-D, le type d'erreurs introduit désynchronise le train-ligne. Nous supposons que la correction est assurée par les techniques de dissimulation mentionnées en IV.2..

c - erreurs de transmission sur les types

Les informations de type de zones (1, 2 ou 3) sont à considérer comme des entêtes de blocs précisant la nature des informations transmises pour cette zone :

zone 1 et 2 : information de longueur de zone = adressage différentiel

zones 3 : nb d'erreurs de prédiction et valeurs de celles-ci.

Une erreur sur le type 3 (transformé après transmission en type 1 ou 2) désynchronise le train-ligne puisque la nature et donc le nombre des informations attendues à la suite de cette entête n'est pas le même. Nous supposons comme ci-dessus que ces erreurs sont corrigées par dissimulation.

Par contre, une interversion de types (1 + 2 ou 2 + 1) est intéressante à étudier ce qui revient à dire que des zones fixes à l'émetteur seront considérées comme compensées en mouvement au récepteur et vice-versa.

d - conclusion

L'une des conclusions expérimentales qui est apparue lors des simulations semble être que la nature et le type de l'injection d'une erreur de transmission est très rapidement sans influence, lors des effets de propagation temporelle de cette erreur. Nous avons procédé ainsi ensuite à des injections d'erreurs fixes en amplitude et en taux mais aléatoires en localisation. Par contre, de la localisation géométrique de l'injection dans l'image (au début ou en fin d'image) dépend l'ampleur de la propagation.

IV.4. - Techniques de correction des erreurs

Au vu de l'observation des effets de propagation temporelle, (voir image en annexe), nous avons recherché à mettre en oeuvre des techniques simples de correction algorithmique visant à atténuer voire annuler les effets de propagation.

Deux séries de tests de méthodes de correction ont été effectuées :

- sans contraintes de débit supplémentaire
- avec la contrainte que le débit supplémentaire du à l'adjonction de techniques de correction représente environ 5 % du débit total (34 Mb).

IV.4.1. - Introduction d'informations PCM

Comme pour les codeurs APCM intraimage, l'une des techniques envisageables pour combattre la propagation des erreurs dues au caractère DPCM du codeur est d'introduire arbitrairement des lignes d'images (ou des pixels) codées en PCM et qui sont plus robustes à la propagation des erreurs.

L'amélioration bien que sensible due à cette technique de correction est peu efficace notamment vis-à-vis de la propagation temporelle des effets d'erreurs. Il convient donc plutôt d'orienter les techniques de correction vers les paramètres intervenant directement dans les modes temporels de prédiction du codeur.

IV.4.2. - Transmission de vecteurs de déplacement

Le codeur étudié (cf. schéma général) fonctionne sur le principe d'une estimation "symétrique" au codeur comme au décodeur des vecteurs \vec{D} de déplacements en chaque point d'image. Ce vecteur recalculé au décodeur n'est donc pas transmis.

Estimé à partir d'informations reconstruites et donc peut-être erronées s'il s'est produit auparavant dans la séquence déjà des erreurs de transmission, le vecteur \vec{D} intervient en effet :

- dans le calcul des prédictions
- dans le choix de stratégies de prédiction entre les différents modes de prédiction
- dans le calcul d'autres estimées de mouvement.

L'ensemble de ces causes combinées crée une propagation temporelle importante. Nous avons donc testé quelle protection apportait la transmission de vecteurs de déplacement (1 vecteur \vec{D} tous les N pixels, N étant choisi selon la contrainte de débit que l'on s'est fixée).

Nous présentons en Annexe les résultats concernant ce type de protection sous la forme des images reconstruites au décodeur après adjonction des schémas de correction ainsi que l'évolution temporelle de la disparité entre l'image initiale et l'image reconstruite.

V - Conclusion

Des résultats présentés en Annexe, nous dégageons les principaux enseignements sui-

DISSIMULATION ET CORRECTION DE LA PROPAGATION TEMPORELLE
D'ERREURS DE TRANSMISSION - JP de la Tribonnière, C. Labit
DISSIMULATION AND CORRECTING TECHNIQUES TO TEMPORAL EFFECTS OF
CHANNEL ERRORS - JP de la Tribonnière, C. Labit

vants :

.) l'introduction d'une erreur isolée sur une image seulement crée un effet local de distorsion qui, lorsque des techniques de correction sont employées, s'évanouit temporairement rapidement en créant un bruit granulaire de distorsion non critique visuellement

.) lorsqu'un flot continu d'erreurs est injecté, cet effet de bruit granulaire s'étend de manière éparse sur toute l'image avec, parfois, des effets d'accumulation ce qui crée des zones erronées non seulement détectables mais gênantes visuellement

.) les techniques de dissimulation sont efficaces et ne créent temporairement tout au long de la séquence qu'un faible bruit granulaire. Les dissimulations intra ou inter-images semblent être équivalentes

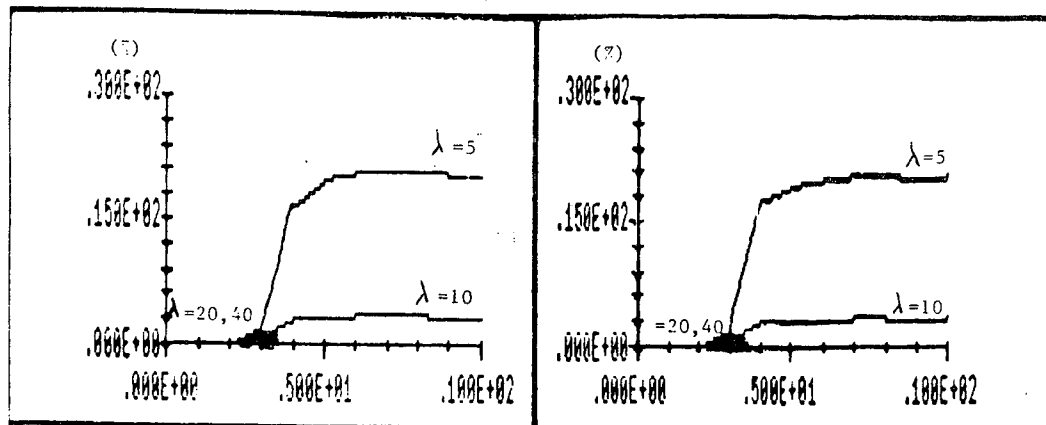
.) parmi les techniques de correction testées, la plus efficace est de loin celle qui protège les vecteurs estimés de déplacement. Sans atteindre un évanouissement complet de la propagation temporelle, nous atteignons un régime quasiment stable dont la valeur asymptotique est plus ou moins élevée selon la contrainte de débit de correction à laquelle on s'astreint. De plus l'amélioration de l'image est visuellement très sensible lorsque de telles techniques sont employées.

Cette étude sera poursuivie dans le cadre de simulation d'autres schémas de codage multimodes intra/interimage avec notamment la définition de techniques de protection/correction lorsque dans l'algorithme de codage lui-même est prévue la transmission du vecteur de déplacement \vec{D} estimé par blocs.

REFERENCES

- [1] R.J.Arguello, H.R.Sellner, J.A.Stuller : "The effects of Channel errors in the differential pulse-code modulation Transmission of sampled Imagery". IEEE Trans. on Com., vol.19, n°6, déc. 1971.
- [2] D.J. Connor : "Techniques vor reducing the visibility of Transmission errors in digitally encoded video signals". IEEE Trans. on Com., vol.21, n°6, June 1973.
- [3] M.C.W. Van Buul : "Hybrid D-PCM. A combination of PCM and DPCM". IEEE Trans.1978.
- [4] V. Guionnet, P. Sallio : "Système expérimental de T.V. numérique à 52Mb/s. Evaluation subjective de l'effet des erreurs de transmission". Revue de radiodiffusion-Télévision, n°54, 1978.
- [5] N.F. Maxemshuk, J.A. Stuller : "Reduction of Transmission Error propagation in adaptively predicted DPCM encoded pictures". BSTJ, vol.58, n°6, July-Aug. 1979.
- [6] P. Stammitz : "Error protection of 34Mb/s DPCM encoded TV signal with multiple error correcting BCH codes". Intern. Zürich Seminar on Digital Communications, mars 1980.
- [7] D. Barba, J.Dewitte : "Reduction of effects of source noise and transmission errors in adaptive DPCM". PCS 1981, Montréal.
- [8] K.N.Ngan, R. Steele : "Enhancement of PCM and DPCM Images corrupted by transmission errors". IEEE Trans. on Com., vol.30, n°1, Janvier 1982.
- [9] F.Kretz : "Codage MICD à prédiction adaptative en T.V. : qualité visuelle, réduction de débit et susceptibilité aux erreurs de transmission". Annales des Télécom., tome 37, n°7-8, Juillet-Août 1982.
- [10] F.J. Garcia-Ugalde : "Protection contre les erreurs pour la transmission de signaux de télévision codés numériquement". Thèse Doctorat Ingénieur, Univ. de Rennes I, Déc. 1982.
- [11] W.W. Peterson, E.J. Weldon : "Error correcting codes". MIT Press, 1972.
- [12] K.Janac, N.J.A. Sloane : "The detection of long error bursts during transmission of video signals". BSTJ, vol.61, n°8, Oct. 1982.
- [13] P.Robert : "Définition d'un schéma de codage multimodes, avec compensation de mouvement, pour les séquences d'images de télévision". Thèse Doctorat-Ingénieur Univ. de Rennes I, Nov. 1983.
- [14] C.Labit, P.Robert, J.Dewitte : "Optimization of addressing bit rate in a multimode coder with motion compensation for broadcast television". Intern. Zürich Seminar, Fev. 1984.

DISSEMINATION ET CORRECTION DE LA PROPAGATION TEMPORIELLE
D'ERREURS DE TRANSMISSION - CP de la Tricommère, J. Lacroix
DISSEMINATION AND CORRECTION TECHNIQUES OF TEMPORAL EFFECTS OF
CHANNEL ERRORS - CP de la Tricommère, J. Lacroix



Comparaison des techniques de dissimulation interimage (courbe de gauche) et intrimage (courbe de droite). Courbes des différences seuillées (pour plusieurs valeurs du seuil λ) entre les images (de n°3 à n°10) reconstruites avant et après dissimulations.



Image n°3 originale (image où sont injectées les lignes dissimulées)

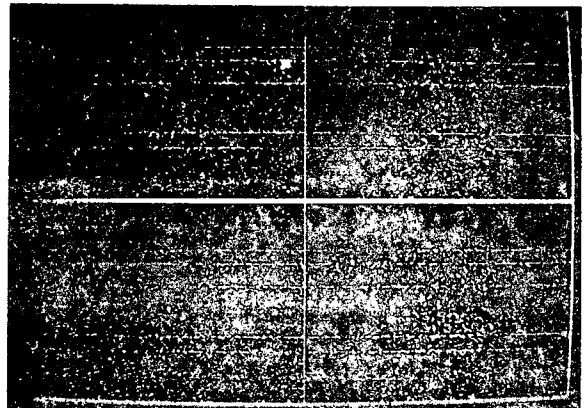


Image des différences seuillées (sur un quart de l'image) (seuil $\lambda=5$) pour 4 images (n°3,4,5,6) dans le cas de la dissimulation intrimage.

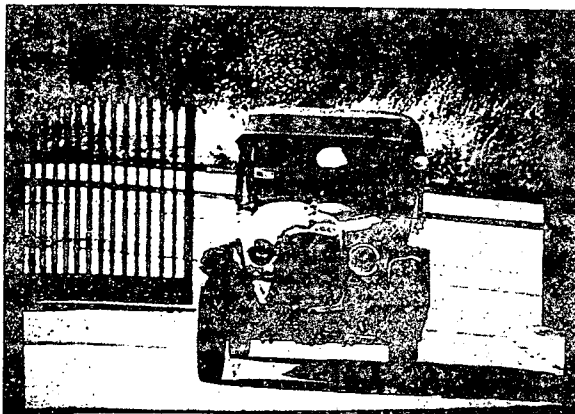


Image reconstruite (n°3) après dissimul. interimage

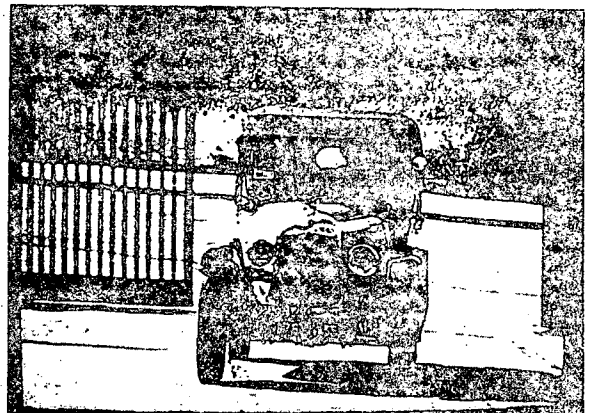
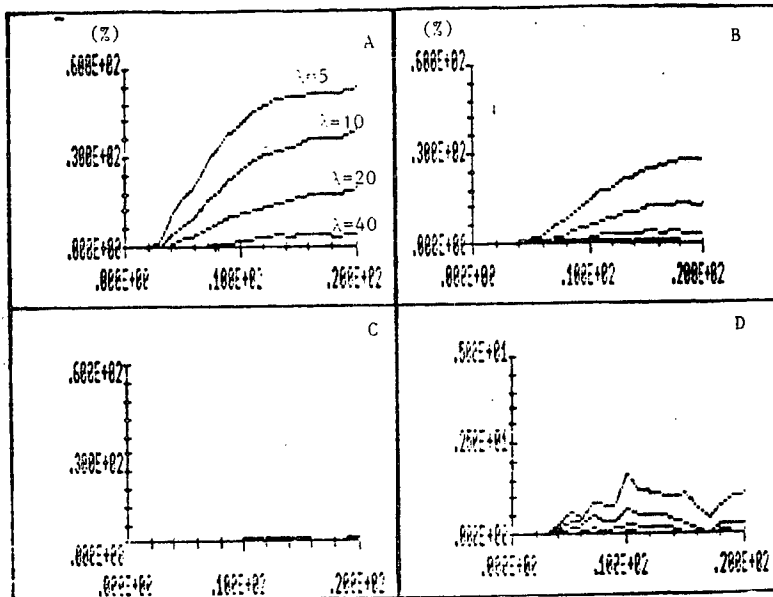


Image reconstruite (n°3) après dissimul. intrimage

CONSERVATION ET TRANSMISSION DE LA PRÉPARATION TEMPORIELLE
D'ERREURS DE TRANSMISSION - Cf. de la Tribonnière, P. Lacroix
TECHNIQUES D'INJECTION D'ERREURS TECHNIQUES ET TEMPORAL EFFET OF
INJECTION OF ERRORS - Cf. de la Tribonnière, P. Lacroix



Courbes des différences seuillées (pour 4 valeurs de seuils $\lambda=5,10,20,40$) entre les images reconstruites au codeur et au décodeur :

A : sans correction

B : avec correction par D et contrainte de débit (5%)

C : avec correction par D (tous les 10 pixels)

D : id. que C mais grossissement d'échelle
Injection d'erreurs sur les images n°3 à n°10 et observation des images n°3 à n°20.



Image n°4 sans correction des erreurs



Image n°4 avec correction par D et contrainte de débit (5%) .

Disposition :

A	B
C	D

A : image originale

B : image reconstruite au codeur

C : image reconstruite au décodeur après "traitement" des erreurs

D : image des différences seuillées ($\lambda=10$) B-C



Image n°10 sans correction des erreurs

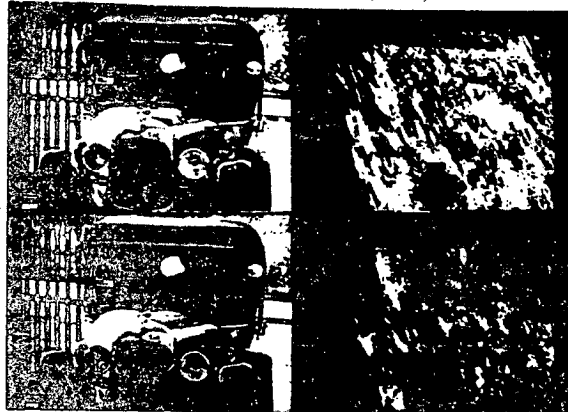


Image n°10 avec correction par D (5%)

DISSEMINATION ET CORRECTION DE LA PROPAGATION TEMPORALE
 L'EFFET DE TRANSMISSION - 37 de la Tribunaire, 2. Labit
 DISSEMINATION AND CORRECTION TECHNIQUES OF TEMPORAL EFFECTS OF
 CHANNEL SPREAD - 37 de la Tribunaire, 2. Labit

Résultats sur l'image n°20

image reconstruite au image des diff. seuillées
 décodeur ($\lambda=10$)

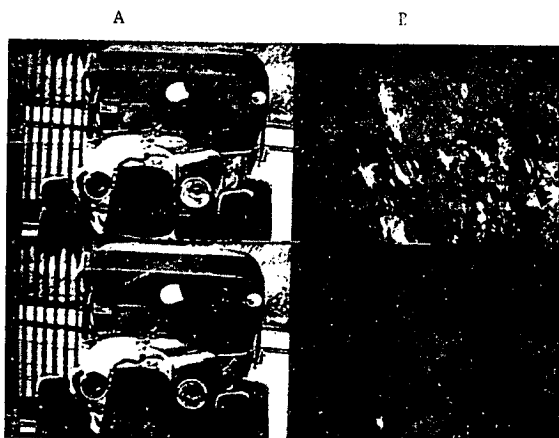


A

B

sans correction

avec
 correction
 par
 (5%)



A

B

avec correction par D
 (tous les 10 pixels).

Disposition :

A | B

A : image reconstruite au décodeur

B : image des différences seuillées ($\lambda=10$)

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique

